The Stability of Incremental Analysis Update

Lawrence L. Takacs *

Science Systems and Applications Inc., Lanham, Maryland

Max J. Suárez

University Space Research Association, Columbia, Maryland

Ricardo Todling

Global Modeling and Assimilation Office,

NASA Goddard Space Flight Center, Greenbelt, Maryland

Corresponding author address: Lawrence L. Takacs, Science Systems and Applications Inc., Lanham, Maryland

E-mail: Lawrence.L.Takacs@nasa.gov

ABSTRACT

A recent attempt to downscale the 50 km MERRA-2 analyses to 7 km revealed an instability associated with the Incremental Analysis Update (IAU) procedure that has thus far gone unnoticed. A theoretical study based on a simple damped harmonic oscillator with complex frequency provides the framework to diagnose the problem and suggests means to avoid it. Three possible approaches to avoid the instability are to: (i) choose an "ideal" ratio of the lengths of the Predictor and Corrector steps of IAU based on a theoretical stability diagram; (ii) time average the background fields used to construct the IAU tendencies with given frequency; or (iii) apply a digital filter modulation to the IAU tendencies. All these are shown to control the instability for a wide range of resolutions when doing up- or down-scaling, experiments with the NASA/GMAO atmospheric general circulation model. Furthermore, it is found that combining IAU with the ensemble re-centering step typical of hybrid ensemble-variational approaches, also results in an instability based on the same mechanisms in the members of the ensemble. An example of such occurrence arises in an experiment performed with the GMAO 12.8 km hybrid 4D-EnVar system. Modulation of the ensemble IAU tendencies with a digital filter is shown to avoid the instability. In addition, the stability of certain 4DIAU implementations is analyzed and a suggestion is made to improve its results, though a complete study of this subject is postponed to a follow up work.

1. Introduction

The Incremental Analysis Update (IAU) was introduced by Bloom et al. (1991, 1996) as an initialization procedure for three-dimensional atmospheric data-assimilation systems (DAS). Since then it has been used successfully at several centers and in multiple applications (e.g., Schubert et al. 1993, Lorenc et al. 2000, Lee et al. 2006, Ourmières et al. 2006, Balmaseda et al. 2008, Carton and Giese 2008), and a four-dimensional extension of the idea has been developed by Lorenc et al. (2015), and used elsewhere (e.g., Buehner et al. 2015; Lei and Whitaker 2016).

For practical reasons, most centers doing numerical weather prediction (NWP) and reanalyses of historical observations maintain some means of reproducing the lengthy calculations of a full data-assimilation system without having to rerun the expensive analysis software, but by only rerunning the model. This can be useful if additional diagnostics of the assimilated state are required that were not included in the original run, or to test, inexpensively, how model changes affect forecasts, albeit without feedback from the analysis. At NASA's Global Modeling and Assimilation Office (GMAO) the means of running the model from pre-existing analyses is referred to as "Replay". The GMAO Replay is fundamentally dependent on IAU-based strategies used for initialization.

When the Replay strategy was initially applied by Putman et al. (2017; pers. comm.) in an attempt to downscale the MERRA-2 Reanalysis (Gelaro and coauthors 2017) to high-resolution, the model was found to become unstable after a few weeks of integration. Replaying with low-resolution versions of the model showed no symptoms of the instability.

More recently, a candidate upgrade running parallel to the GMAO 12.8 km hybrid 4D-EnVar system (Todling and El Akkraoui 2018) was found to show signs of a developing instability in the members of its underlying 50 km ensemble. Among the changes being evaluated in the parallel system was the re-introduction of the procedure typically referred to as "ensemble recentering",

which simply replaces the analysis ensemble mean with the hybrid analysis. The preceding GMAO 25.6 km hybrid 3D-Var system recentered the members of its 100 km ensemble with no detrimental impact. Preliminary evaluations of the upgrade of this system to the 12.8 km hybrid 4D-EnVar system, however, had shown instabilities when recentering was applied to the members of the ensemble. Without a solution to the problem, the January 2017 initial release of the GMAO 12.8 km hybrid 4D-EnVar system did not include a recentering procedure. When trying to reinstate recentering in the latest candidate upgrade of the 12.8 km system, once again the instability manifested itself. With insight from the work being done to understand the instabilities in the Replay context, it became evident that the instability being experienced in the ensemble part of the cycled hybrid system was an instance of the same problem; indeed, as explained in this work, recentering can be viewed as a form of Replay.

A simple linear mathematical analysis, focused on Replay, explains the problem and provides a means to avoid the instability. In this article, we take a second look at the work of Bloom et al. (1996) with the purpose of understanding the instability associated with IAU, and how it impacts Replay strategies in general. We also provide a brief insight on the implications of the results here to extensions of IAU to four-dimensional data assimilation.

The next section provides details on both IAU and Replay, followed, in Sec. 3, by a stability analysis of these two mechanisms. Section 4 reports on the manifestation of the instability in IAU/Replay experiments with the GMAO Atmospheric General Circulation Model (AGCM). Section 5 presents the relationship between model resolution and the development of instabilities associated with IAU and Replay. Section 6 presents three possible ways of avoiding the instability. Section 7 reports on the manifestation of the instability in the ensemble supporting the GMAO hybrid 4D-EnVar data assimilation system, how it relates to the combination of IAU and recentering, and it shows how the instability can be removed. Section 8 provides some insight on how to

re-configure current 4DIAU formulations to possibly improve results from hybrid 4D assimilation systems. Some closing remarks and future work appear in Sec. 9.

2. Overview of IAU and Replay

Independently of balance constraints imposed on the underlying analysis system, differences in resolution and intrinsic model balances require some type of initialization to be implemented to minimize the effect of spurious waves from affecting how a model adjusts to changes imposed by the analysis (e.g., Kalnay 2003, Sec. 5.7). The IAU of Bloom et al. (1996) is an alternative to nudging which, in principle, has better model responses than nudging itself.

Figure 1(a) provides a schematic illustration of a 3-hourly, first-guess at the appropriate time (FGAT; Lawless 2010, Massart et al. 2010), implementation of IAU. For example, for the 00Z analysis, the cycle begins at 21Z — the end of the initialized forecast of the previous assimilation cycle. A free-running (00Z *Predictor*) model is integrated for six hours, from 21Z to 03Z, generating "background" fields at 21Z, 00Z, and 03Z (cross-hatched green arrow and red circles). The FGAT 3D analysis combines these backgrounds with observations to produce an analyzed state valid at 00Z. The model is then rewound to 21Z and run a second time, for another six hours, starting from the same initial and boundary conditions used for the 21Z Predictor step, but now forced with a constant tendency term built as the difference between the analysis and the background at 00Z divided by a timescale τ_{iau}, normally taken as six hours (shaded triangle). This second model integration is referred to as the *Corrector* step for the 00Z analysis. Note that, with IAU, the concatenation of the Corrector segments (solid green arrows) represent a continuous integration of the model in which the effects of the analysis appear as an extra diabatic forcing refreshed six-hourly.

The Replay mechanism is similar to IAU, but uses pre-existing analyses to build IAU tendencies. The Replay integration is thus a blend of given analyses and model results. If the model is identi-

cal to the one used to produce the analysis, the results of the Corrector step are identical to those from the original assimilation. If the analysis increments are saved along the way during the original assimilation cycle, a Replay with an identical model can simply be done using the available increments — bypassing the Predictor step of Fig. 1(a). In the general purpose implementation of Replay, when the model is allowed to differ from the original model used in the assimilation cycle one would follow the procedure shown in Fig. 1(b). This form of Replay applies to cases when the model changes configuration (e.g., resolution), or when any of its internal procedures (e.g., physics, parameterizations, etc.) change for various reasons. These cases require the Predictor step (cross-hatched green arrow and red circle), and calculation of a Replay increment built as the difference between a pre-existing analysis and the new background, both valid at the same time. Results from this form of Replay cycle (solid green arrows) do not reproduce those of the original assimilation cycle, but model-generated products are still closely related to those of the fully cycled assimilation.

For illustration purposes, Fig. 2 compares the power spectrum amplitudes of the rate of change of surface pressure due to dynamics in (hPa/day), averaged over the tropics from 30-South to 30-North, for the three cases defined by 1) direct insertion of analysis increments in a non-IAU-based DAS (red curve), 2) the use of IAU in a DAS (black curve), and 3) a free-running model with no data assimilation (blue curve). The long-period peaks are physical and correspond to tidal effects at diurnal (24 hours), semi-diurnal (12 hours) and terdiurnal (8 hours) periods. We also see a peak at 6 hours corresponding to the frequency of the analysis update in the DAS cases. Direct insertion shows considerably larger power in the 6-hour mode as compared to IAU. While the AGCM is generally capable of damping the higher-order harmonics right after data insertion, the repeated forcing from direct data insertion during an extended assimilation results in an enhanced excitation

of higher-order harmonics when compared to the IAU procedure. Fig. 2 clearly illustrates the benefits of IAU for damping these higher frequency modes.

3. Simplified Analysis of the Stability of IAU/Replay

The analysis of Bloom et al. (1996) compares different ways of inserting the analysis increment into a linear oscillator with complex frequency. They compared IAU to direct insertion of the analysis, which they called *intermittent update*. They showed that, with IAU, a neutral oscillator's response to the insertion has decreasing amplitude with increasing frequency; while with direct insertion, the response's amplitude is independent of frequency. This is the "initialization" purpose of IAU, to minimize excitation of the high frequency modes by the analysis increment. This, however, is not an indication of the stability of the procedure.

We begin by briefly reminding the reader of the Bloom et al. (1996) frequency analysis. What follows uses their idealized model and provides the grounds for the considerations relevant to the present work. IAU is modeled as a damped oscillator, with complex frequency $\tilde{\omega} = \omega + i\kappa$, driven by a piece-wise constant forcing over the analysis interval:

$$\frac{dU}{dt} = i\tilde{\omega}U + \frac{\Delta U}{\tau_{iau}}, \quad t_n < t < t_n + \Delta t_C, \tag{1}$$

where ΔU is the analysis increment; κ represents physical or numerical damping in the model, which will be critical for understanding the IAU stability; Δt_C is the length of the Corrector step (here also defined as the analysis interval); and τ_{iau} is a disposable constant parameter. Integrating (1) from t_n to t gives:

$$U(t) = U(t_n)e^{i\tilde{\omega}(t-t_n)} - \frac{i}{\tilde{\omega}}\frac{\Delta U}{\tau_{iau}}\left[e^{i\tilde{\omega}(t-t_n)} - 1\right], \tag{2}$$

and evaluating this expression at the end of the interval we get, after some rearrangement:

$$U(t_n + \Delta t_C) = U(t_n)e^{i\tilde{\omega}\Delta t_C} + \Delta U\, e^{i(\tilde{\omega}\Delta t_C/2)}\left(\frac{2\sin(\tilde{\omega}\Delta t_C/2)}{\tau_{iau}\tilde{\omega}}\right). \tag{3}$$

Bloom et al. (1996) compares (3) to the solution using direct insertion, when the analysis increment is added instantaneously at $t_n + \frac{\Delta t_C}{2}$. In this case, $U(t_n)$ is propagated for half the interval, the increment is added, and the sum is propagated for the second half of the interval:

$$\begin{aligned} U(t_n + \Delta t_C) &= \left(U(t_n) e^{i(\tilde{\omega}\Delta t_C/2)} + \Delta U \right) e^{i(\tilde{\omega}\Delta t_C/2)}, \\ &= U(t_n) e^{i\tilde{\omega}\Delta t_C} + \Delta U e^{i(\tilde{\omega}\Delta t_C/2)}. \end{aligned} \tag{4}$$

In both (3) and (4), the first term represents the free propagation of the initial conditions from times t_n to $t_n + \Delta t_C$. The second term is the effect of the increment. In (4), the increment is simply propagated from $t_n + \frac{\Delta t_C}{2}$ to $t_n + \Delta t_C$. In (3), the increment is similarly propagated, but now modulated by an $\tilde{\omega}$-dependent factor. As discussed in Bloom et al. (1996), this factor ameliorates the impact of the increments on the high-frequency modes — the effect sought from IAU: that is, it acts as a low-pass filter. The Bloom et al. (1996) study, however, provides no evaluation of the stability of IAU.

At this point, we might be wary of two facts: (i) the mathematical arguments presented so far make no reference for how the analysis increment is calculated, which might be relevant to stability considerations; and (ii) although the amplitude of the response to the increment is nicely handled at higher frequencies by IAU, for a sufficiently inviscid model ($\kappa \to 0$), its sense reverses wherever $\sin(\omega \Delta t_C/2)$ changes sign, allowing it to be amplifying, rather than restoring, at some frequencies.

In 3D-Var schemes, such as used in MERRA-2, the analysis is produced at the single synoptic time within the IAU cycle (the Corrector step). The analysis increment can then be written as

$$\Delta U = U_A(t_n + \Delta t_P) - U_B(t_n + \Delta t_P), \tag{5}$$

where, U_B and U_A are the background and analysis states, respectively, and Δt_P is the length of the free forecast made to produce the background (the Predictor step). To include both DAS and Replay situations, we write the "analysis" in (5) as a linear combination of the background, U_B,

and a "target" state, U_T, which can represent either the observations when considering the regular DAS mode, or a prescribed state when the Replay mode is considered instead. That is,

$$U_A(t_n + \Delta t_P) = K U_T(t_n + \Delta t_P) + (1 - K) U_B(t_n + \Delta t_P)\,, \tag{6}$$

where K represents the analysis *gain*, which weighs the background and the target states. Note that if we draw to the target perfectly ($K = 1$), the background has no influence on U_A. This is the methodology used in Replay mode, i.e., the model is forced by a *pre-existing* analysis which cannot be influenced by the current background. In the case of a full IAU-based DAS cycle the analysis is a linear combination of the model background and the observations ($K < 1$). Combining (6) and (5), the increment becomes

$$\Delta U = K\left[U_T(t_n + \Delta t_P) - U_B(t_n + \Delta t_P)\right]. \tag{7}$$

The discussion here will assume that both U_T and U_B satisfy the homogeneous form of (1), and that U_B matches the full solution, U, at $t = t_n$, so that (7) becomes

$$\Delta U = K\left[U_T(t_n) - U(t_n)\right] e^{i\tilde{\omega}\Delta t_P}\,, \tag{8}$$

and (3) can be rewritten as

$$U(t_n + \Delta t_C) = U(t_n) e^{i\tilde{\omega}\Delta t_C} + \frac{\Delta t_C}{\tau_{iau}} K\left[U_T(t_n) - U(t_n)\right] e^{i\tilde{\omega}(\Delta t_P + \Delta t_C/2)} \mathrm{sinc}\left(\frac{\tilde{\omega}\Delta t_C}{2}\right), \tag{9}$$

where $\mathrm{sinc}(x) = \sin(x)/x$.

Since we are assuming U_T satisfies the homogeneous equation, we can subtract $U_T(t_n + \Delta t_C) = U_T(t_n) e^{i\tilde{\omega}\Delta t_C}$ from (9) and, defining $\delta U = U - U_T$, obtain

$$\delta U(t_n + \Delta t_C) = \delta U(t_n) e^{i\tilde{\omega}\Delta t_C} \times \left[1 - \frac{\Delta t_C}{\tau_{iau}} K\, e^{i\tilde{\omega}(\Delta t_P - \Delta t_C/2)} \mathrm{sinc}\left(\frac{\tilde{\omega}\Delta t_C}{2}\right)\right]. \tag{10}$$

The amplification factor for each IAU cycle is the magnitude of the quantity in the square brackets. Note that this amplification factor is associated with the ratio of the *errors* in the solution over one

analysis cycle. In Replay mode ($K = 1$), with $\Delta t_P = \Delta t_C/2$ and $\tau_{iau} = \Delta t_C$ (which are standard IAU choices) it takes the simple form

$$\frac{|\delta U(t_n + \Delta t_C)|}{|\delta U(t_n)|} = e^{-\kappa \Delta t_C} \left| 1 - \mathrm{sinc}\left(\frac{\tilde{\omega} \Delta t_C}{2} \right) \right|. \tag{11}$$

A plot of (11) is shown in Fig. 3 for $\Delta t_C = 6$ hours. For low frequency oscillations of δU (periods much longer than Δt_C), δU is strongly damped to zero, and the solution quickly tracks the target solution, which is the desired behavior of IAU. At high frequencies, the amplification factor approaches one, which is also the desired behavior of the IAU, i.e., eliminating high-frequency responses to the increment and leaving U unmodified. However, for frequencies corresponding to periods between Δt_C and $\Delta t_C/2$, the amplification factor can exceed one, unless the oscillation is sufficiently suppressed by the model's natural damping. Note that the amplification peaks at a period slightly longer than $\frac{2}{3}\Delta t_C$. For $\Delta t_C = 6$ hours and $\kappa = 0$, it peaks at a period of 4.2 hours, with an e-folding time of approximately 30 hours.

A cross-section at $\kappa = 0$ of the Replay mode ($K = 1$) of Fig. 3 is shown in Fig. 4 (solid black curve) together with a "typical" DAS mode case ($K = 0.5$; dashed red curve). In this case, the DAS mode presents reduced amplitude of the unstable modes, indicating improving stability. In fact, in an actual data assimilation setting, the emergence of such a mode would result in increasing discrepancies between the observations and the background for modes associated with the instability, leading to a lower effective K and further reducing the amplification. This must account for the evidence that IAU is more stable in DAS mode than in Replay mode.

A heuristic argument for the instability is presented in Fig. 5. The idea is that the fixed increment used in IAU and Replay is based on the value of the background at a specific time in the interval, in this case ($\Delta t_P = \Delta t_C/2$) the mid-point, and that since the increment is applied uniformly over

the interval, its instantaneous effect can force the solution both toward or away from the target solution.

The argument for instability is that for oscillations with periods shorter that Δt_C the increment can, on average, drive the solution away from the target trajectory. The schematic in Fig.5 illustrates this for an oscillation with a period of $\frac{2}{3}\Delta t_C$, which is close to the most unstable case. The increment computed at the mid-point has a restoring effect only over the central third of the interval, while driving the solution away from the target over the first and last thirds. Similar reasoning applied to periods of Δt_C or $\frac{1}{2}\Delta t_C$ (not shown) argues for these oscillations to be neutral, since they result in exact compensation of amplifying and restoring effects over the full interval, and that periods longer than the interval are dominated by restoring influences. This is exactly the behavior shown in Fig 3, where for $\kappa \to 0$, the first band of unstable oscillations occurs between these periods, and that periods longer than Δt_C are absolutely stable. Other unstable bands occur at even shorter periods, but because of the behavior of the sinc function, these have weaker amplification and are not of practical importance.

4. Manifestation of the Instability in Replay Experiments with a GCM

Figure 6 shows a comparison between a high-resolution (cubed-sphere C720, or an average grid size of $\sim$12.8 km) assimilation using IAU (left panel), and a lower-resolution (C360 or $\sim$25.6 km) Replay to this assimilation using the same model (right panel). To highlight the issue, we show the rate of change of surface pressure due to dynamics along the Equator for a 10-day period in 2017. The analysis is dominated by the semi-diurnal tide, which shows no pathological behavior. The Replay tracks the analysis well for a few days, but after a week or so it becomes unphysical.

Figure 7 shows "snapshots" of the surface pressure tendency from dynamics (i.e., not including the analysis or Replay increments) on 27 January 2017, for the same two runs. The DAS shows

the characteristic semi-diurnal tide in the tropics and the propagating synoptic systems in mid-latitudes. The Replay experiment (lower panel) is dramatically different; it is characterized by large-amplitude large-scale features over all regions of the globe that overwhelm the physical features in the DAS.

Using data from the last five days (January 25-29, sampled every 5 minutes) of the C720 DAS and C360 Replay experiments, the tropically-averaged (30S to 30N) Fourier decomposition of the sea-level pressure was computed for periods from 1-hour to 1-day; this is shown in Fig. 8. We see the large-amplitude 12-hour physical mode associated with the semi-diurnal tide, as well as the 8-hour terdiurnal tide, in both experiments. We note that these modes (which are in the low-frequency stable region of Fig. 3) are well reproduced by Replay. In the band with periods between 6 and 3 hours, however, the Replay experiment shows spurious amplification with a peak at approximately 4.2 hours, in agreement with the period of the most unstable mode in Fig. 4. Note that the DAS also has peaks at the 6-, 4.8-, and 4-hour diurnal harmonics, and that the secondary peak in the amplified Replay response corresponds to the 4.8-hour harmonic.

5. Instability Dependence on Model Horizontal Resolution

As mentioned in the Introduction, low-resolution Replay experiments showed no evidence of this instability. We see from Fig. 3 that the instability is controlled for systems containing a damping time-scale of 30 hours or less. To gain insight into the effective damping time-scales within the GMAO AGCM, the model state in the C360 Replay experiment of Fig. 7 was used as initial conditions for straight model forecasts (no IAU) at cubed-sphere resolutions ranging from C48 ($\sim$ 192 km) to C1440 ($\sim$ 7 km). The most unstable mode was thus allowed to spin down to its normal state. The upper panel in Fig. 9 shows the temporal behavior of the globally-averaged magnitude of the surface pressure tendency due to dynamics for each of the model resolutions

(solid curves). In addition, e-folding damping (dashed) curves are shown with time-scales chosen to approximate the behavior of the model results. The lower panel of Fig. 9 shows the asymptotic behavior ($\approx$ 100 hours) of the effective damping time-scales as a function of horizontal resolution. We see that at very low resolution (C48), the effective damping time scale is 22-hours, which is well within the stable region in Fig. 3. In practice, the C48 Replay runs successfully without ever showing instability issues. As the model resolution increases, the corresponding effective damping time scale also increases. The analysis summarized in Fig. 3 suggests that the C90 Replay should be marginally unstable, and in fact, it runs for well over a month before experiencing problems. All higher resolution runs (C180 and greater), with growth rates well within the unstable region, fail within 1-2 weeks.

This dependence of the model's damping of the mode on the model's resolution resolution is almost entirely due to the dynamics, becoming less dissipative with increasing horizontal resolution. To quantify this, we repeated the spin-down experiments (not shown) with the model's explicitly computed diabatic and frictional effects eliminated, so that the model spins down only due to numerical dissipation. Assuming all damping behaves linearly $(1/\tau_{Full} - 1/\tau_{NumOnly} = 1/\tau_{Diabatic})$, gives a value for $\tau_{Diabatic}$ of $\approx$ 230 hours.

The Replay instability also manifests under different Predictor/Corrector strategies. Assuming *multiple* external analysis within a 6-hour window, a simple alternate Predictor/Corrector approach to the standard 3-hour Predictor 6-hour Corrector used in MERRA-2 might be to compute IAU increments based on the current model background and its associated external analysis (i.e., no Predictor step), and hold the increment constant until the next available analysis. At this point, the IAU increment would be recomputed based on the current analysis and current model background. To demonstrate the impact of both horizontal resolution and Corrector frequency on Replay stability using this approach, three C720 Replay experiments using Predictor durations equal to zero

were run with Corrector durations of 3-, 2-, an 1-hour. The Replayed external analyses were taken from the current GMAO C720 hybrid 4D-EnVar DAS having analyses available hourly. The upper-panel of Fig. 10 shows the harmonic decomposition of the sea-level pressure for the Replay experiments using these 3 strategies, while the lower-panel depicts the amplification factors from (10) for the inviscid, Replay ($K = 1$) case. Not surprisingly the Predictor = 0, Corrector = 3-hour case (P0 C3), whose most unstable mode is near the 4-hour frequency, is unstable at C720 just as previously seen for C180 and C360. As we increase the Corrector frequency to 2-hours, the most unstable mode moves closer to 3-hours. This, too, is unstable in the C720 Replay (P0 C2). Finally, when a 1-hour Corrector duration is used, the C720 model's effective damping is strong enough for the run to survive without issues (not shown). However, when the resolution is increased to C1440, even the 1-hour Corrector duration (P0 C1) becomes unstable in modes with periods near 1.5-hours. In all cases, the unstable mode predicted by (10) is in close agreement with the actual unstable mode which manifested in the AGCM. It should also be noted that for all three cases, the errors associated with the longest time-periods do *not* reduce to zero as they should, but rather, asymptote to $\left(1 - \frac{\Delta t_C}{\tau_{IAU}}\right)^N$, where N is the number of Corrector steps within the 6-hour window.

6. Attempts to Control the Instability

a. Sweet Spot

We have seen from the preceding sections that the IAU/Replay instability manifests itself under a variety of paradigms. Two such examples are: the 3-hr Predictor 6-hr Corrector strategy used for the C360 MERRA-2 Reanalysis; and the 0-hr Predictor 1-hr Corrector paradigm used for the C720 hybrid 4D-EnVar DAS. Thus it becomes useful to gain an overall assessment of the instability as a function of the complete range of Predictor/Corrector configurations. Figure 11 generalizes

the Replay ($K = 1$) inviscid stability diagram by displaying the amplification per τ_{iau} interval as a function of the $\Delta t_P/\Delta t_C$ ratio and normalized frequency (or inverse period). We see that the strategies employed in two GMAO assimilation systems — namely, that in the C360 MERRA-2 Reanalysis, with $\Delta t_P/\Delta t_C$=0.5, and that in the C720 hybrid 4D-EnVar DAS, $\Delta t_P/\Delta t_C$=0.0 — amount to unstable regimes when employed in Replay experiments (the roughly 4-hr and 1.5-hr modes, respectively). However, it is also clear from Fig. 11 that there exists a "sweet spot" in the Predictor-Corrector ratio ($\Delta t_P/\Delta t_C = \frac{1}{4}$) for which the dominant instability is shifted to modes with much higher frequencies. For these periods the model's implicit damping characteristics may be sufficient to eliminate the IAU/Replay instability. The stability curves for the Standard and Sweet-Spot IAU/Replay strategies are depicted in Fig. 12 as curves (a) and (b), respectively. We see that while the offending 4-hr instability has been removed, the long time-period modes remain further away from the target solution.

b. Background Averaging

Another approach to stabilize the system is to base the increment on a time average of the background, $\overline{U_B}$, rather than on the instantaneous midpoint value used in (7). This approach was actually used in Orbe et al. (2017) to prevent the C90 model IAU Replays from developing the instability. We begin looking at this option by analyzing the case of using a simple average of duration 2α centered at the midpoint:

$$\overline{U_B}(t_n + \Delta t_P) = \frac{U(t_n)}{2\alpha} \int_{t_n+\Delta t_P-\alpha}^{t_n+\Delta t_P+\alpha} e^{i\,\tilde{\omega}\,(t'-t_n)}\, dt', \tag{12}$$

where $0 \leq \alpha \leq \Delta t_P$. Integrating, this can be written:

$$\overline{U_B}(t_n + \Delta t_P) = U(t_n)\, e^{i\tilde{\omega}\Delta t_P}\, \mathscr{B}(\alpha), \tag{13}$$

where

$$\mathscr{B}(\alpha) = \frac{\sin \tilde{\omega}\alpha}{\tilde{\omega}\alpha}. \tag{14}$$

Substituting this in place of $U_B(t_n + \Delta t_P)$ in (7) and re-arranging, we have:

$$\Delta U = K e^{i\tilde{\omega}\Delta t_P} \left[U_T(t_n) - \mathscr{B}(\alpha) U(t_n) \right]. \tag{15}$$

Now, since U_T times a constant also satisfies the homogeneous form of (1),

$$\delta U_\alpha = U - \mathscr{B}^{-1}(\alpha) U_T \tag{16}$$

satisfies (1) with ΔU given by (15). Its solution can be obtained directly from (10), replacing δU with δU_α:

$$\delta U_\alpha(t_n + \Delta t_C) = \delta U_\alpha(t_n) e^{i\tilde{\omega}\Delta t_C} \times \left[1 - \frac{\Delta t_C}{\tau_{iau}} K e^{i\tilde{\omega}(\Delta t_P - \Delta t_C/2)} \mathrm{sinc}\left(\frac{\tilde{\omega}\Delta t_C}{2} \right) \mathrm{sinc}\left(\tilde{\omega}\alpha \right) \right]. \tag{17}$$

Note that if the average is taken over the entire interval ($\alpha = \frac{1}{2}\Delta t_C$) the system is stable for all frequencies when $\Delta t_P = \Delta t_C/2$. For $\alpha = \frac{1}{3}\Delta t_C$, which is the choice made for the MERRA-2 downscaling and in the work of Orbe et al. (2017), the main unstable maximum corresponds to an amplification of about 3% per replay cycle, compared to about 22% for the case without averaging [see curve (c) in Fig. 12]. For stable solutions, $\delta U_\alpha \to 0$ as the Replay proceeds for many cycles, implying that $U \to U_T/\mathscr{B}(\alpha)$, so it is *amplified* compared to the target solution. Its phase approaches that of U_T for periods longer than the Replay interval, while for shorter periods it can be either in phase or 180 degrees out of phase with U_T. We see in Fig. 12 that the stability characteristics for "Background Averaging" follow closely the original Standard curve until it levels off at a neutral solution. We must remember, however, that its target solution is no longer the original target, but rather, one in which its solution has been amplified.

c. Digital Filter

In Bloom et al. (1996), a generalized version of IAU was considered in which the increment is not added uniformly over the Corrector interval, but is weighted towards the center of the interval. After experimentation with various weighting functions, it was decided that localization of the increment was not justified. In light of the unstable behavior, however, we will revisit this decision. While still requiring that the analysis increment ΔU be held constant over the Corrector interval, we simply iterate (2) for N steps with the introduction of time-dependent weights $g(t)$. Doing so we find:

$$U(t_n+\Delta t_C) = U(t_n)e^{i\tilde{\omega}\Delta t_C} + \frac{\Delta U}{\tau_{iau}} e^{i(\tilde{\omega}\Delta t/2)} \frac{2\sin(\tilde{\omega}\Delta t/2)}{\tilde{\omega}} \sum_{j=0}^{N-1} g_{(N-1)-j}\, e^{ij\tilde{\omega}\Delta t}. \tag{18}$$

Here $N = \Delta t_c/\Delta t$ and Δt is the model time-step. Note that for $g(t) \equiv 1$, (18) reduces to (3) when using the relation:

$$\sum_{j=0}^{N-1} e^{ij\tilde{\omega}\Delta t} = \frac{\sin(\tilde{\omega}N\Delta t/2)}{\sin(\tilde{\omega}\Delta t/2)} e^{i\tilde{\omega}(N-1)\Delta t/2}. \tag{19}$$

Again using the definition $\delta U = U - U_T$, we obtain:

$$\begin{aligned} \delta U(t_n+\Delta t_C) &= \delta U(t_n)e^{i\tilde{\omega}\Delta t_C} \\ &\times \left[1 - Ke^{i\tilde{\omega}(\Delta t_P+\Delta t/2-\Delta t_C)} \frac{\Delta t}{\tau_{iau}} \mathrm{sinc}\left(\frac{\tilde{\omega}\Delta t}{2}\right)\right. \\ &\left. \times \sum_{j=0}^{N-1} g_{(N-1)-j}\, e^{ij\tilde{\omega}\Delta t}\right]. \end{aligned} \tag{20}$$

Polavarapu et al. (2004) pointed out that in the 3D-Var case the use of the time-dependent weights is equivalent to applying a digital time filter to the analysis increment using constant weights; we will be considering time-dependent IAU schemes with the weights they proposed which were taken from Lynch and Huang (1992) and, for completeness, reproduced here:

$$g_n = \frac{\sin\left(\frac{2\pi n}{N}\right)}{\pi n} \mathrm{sinc}\left(\frac{2\pi n}{N+1}\right). \tag{21}$$

Curve (d) in Fig. 12 shows the effect of applying the digital filter (DF) to the standard 3-hour Predictor, 6-hour Corrector case. We see now a dramatic improvement in the stability characteristics as well as an improved response in modes with long time periods.

d. Experiment Results

To test which of the above three methods is "best" at Replaying to a target analysis, a control AGCM experiment was run at C360 resolution for 2 months (initialized from MERRA-2 on 1 November 2016) from which model output was generated every 30-minutes. Using the same model, in identical-twin settings, additional experiments were run (initialized from MERRA-2 on 20 November 2016) that "Replayed" to the control experiment output using each of the three different IAU configurations discussed above, i.e., a) "Sweet Spot" using a Predictor step of 1.5-hours together with a Corrector step of 6-hours, b) "Background Averaging" using a Predictor step of 5-hours (in which a 4-hour time-averaged background state was computed from 1-hour to 5-hours, centered around the synoptic time of 3-hours) together with a Corrector step of 6-hours, and c) "Digital Filter" using a Predictor step of 3-hours together with a Corrector step of 6-hours during which the analysis increment is time-modulated. Note that if we had started with the same initial condition as that of the control experiment, a standard Replay (3-hr Predictor, 6-hr Corrector) would reproduce the control run perfectly since the analysis increments would remain identically zero; changing the initial condition allows for a definitive test of the three methods to see which one best reproduces the control. Five-day forecasts, initialized at 00z for each of the 31 days in December, from each of the Replay experiments were used to assess skills. All forecasts were verified against the original control experiment. Figure 13 shows the 31-member averages of the Northern Hemisphere Extratropics 500-mb geopotential height anomaly correlations. The upper panel shows the 31-sample averaged skill for each experiment, while the lower panel shows the

magnitude of the difference with respect to the control case using the digital filter. The colored boxes depict the magnitude of difference associated with 90% statistical significance. We see a clear statistically-significant advantage to use the DF as compared with the other techniques. It is interesting to note that the Sweet-Spot method performed better than Background Averaging even though the stability diagram appears to show improved characteristics in the longer time-periods. One possible explanation is that while the Background Averaging stability curve is well constrained, the "target" analysis in its measure of error is associated with modes which are amplified compared to the control verification.

7. Manifestation of the instability in a recentered ensemble DAS

As mentioned in the Introduction, we have stumbled on the instability in the cycle of our hybrid DAS in a very peculiar way. In January 2017, GMAO upgraded its hybrid 3D-Var system with a hybrid 4D-EnVar strategy. The upgrade also included an increase in resolution of the forecasting model used in the deterministic hybrid DAS from roughly 25.6 km (C360) to 12.8 km (C720), and an increase in the 32 members of the ensemble from roughly 100 km (C90) to 50 km (C180). While the hybrid 4D analysis is based on a hybrid extension of the Gridpoint Statistical Interpolation analysis (e.g., Kleist et al. 2009; Kleist (2012), and D. F. Parrish, pers. comm.), the ensemble DAS is based on a single-time state update with analyses generated by the Whitaker et al. (2008) ensemble square-root filter (EnSRF) at the center-time of the assimilation window. Accordingly, the deterministic DAS uses a particular flavor of 4DIAU, whereas the ensemble uses the traditional, Bloom et al. (1996), (3D) IAU[1]. That is, whereas the ensemble is analogous to MERRA-2 in its use of IAU, with a 3-hour Predictor step, ($\Delta t_P = 3$ hr), combined with a 6-hour Corrector step, ($\Delta t_C = 6$ hr), the C720 hybrid 4D-EnVar DAS flavor of 4DIAU entails zero-hour Predictor steps, $\Delta t_P = 0$,

[1]The ensemble here does not use 4DIAU as in the pure-ensemble, non-hybrid, exercise of Lei and Whitaker (2016).

combined with hourly Corrector steps, $\Delta t_C = 1$ hr. This particular configuration of 4DIAU differs from that of Lorenc et al. (2015) in that it calculates the hourly, piecewise constant, tendencies by scaling the difference between the 4D analyses and the integrating model state at the corresponding time, and is thus referred to as nudged-4DIAU. The details of the GMAO hybrid system appear in Todling and El Akkraoui (2018).

Early experimentation with the 12.8 km hybrid DAS encountered, what appeared to be at the time, problems with ensemble recentering. Since recentering does not affect ensemble spread, and the latter is what matters for a hybrid analysis, the first version of our 12.8 km system was released without recentering. Although no instabilities ever plagued the first release of the forward processing (FP; real-time), system, we find it discomforting to see non-negligible differences between the ensemble mean and hybrid analyses (not shown). Therefore, in a more recent upgrade being prepared to replace the FP system with various unrelated enhancements, recentering was reinstated. The candidate system (FPP) was run in parallel to the FP system for many months without clear evidence of trouble. However, the subtlety of trouble becomes apparent when the number of surface pressure observations accepted in the ensemble, EnSRF, is compared between the FP system and the candidate parallel experiment. A time series of these, covering over six months of assimilation, is shown in Fig. 14: the black curve is for FP; the blue curve is for the parallel system. The steady decline in the surface pressure observation count is very clear in the parallel system. Interestingly enough, the count of observations of other types is barely affected (not shown). Indeed, such decline in surface pressure observations should lead to a consequent decline in accepted satellite radiance observations. However, the configuration of the EnSRF in the GMAO system mimics closely that of the NCEP system which, for computational reasons, amounts to normally having the ensemble take very few radiance observations to begin with, and thus not be much of a factor here.

Once the steady drop in observation count was detected, we made the decision to turn recentering off. This instant is indicated in Fig. 14 by the left-most vertical (pink) line. Immediately, we see the system recovering with a consequent increase in the count of surface pressure observations. A comparison of the 6-hour surface pressure tendencies from the hybrid deterministic DAS with that from the ensemble mean DAS, taken within cycles 500 to 600, appears in Fig. 15. The patterns here are very similar to the patterns seen in Fig. 7 for the Replay case. In the case here, we see the hybrid DAS tendency (top panel, Fig. 15) is rather comparable to what appears in the top panel of Fig. 7. Similarly, the instability appearing in the ensemble has much the same signature as what appears in the bottom panel of Fig. 7. As we see in Fig. 15, the instability detected in the ensemble goes basically unnoticed in the hybrid DAS. In the typical two-way feedback, the ensemble provides covariance information to the hybrid DAS; this is formed from ensemble perturbations — deviations from the ensemble mean — and are thus not very affected by the instability. The second part of the two-way feedback is established by the hybrid DAS providing its analysis for recentering of the ensemble members.

As it turns out, recentering can be seen as a form of Replay. In the scalar, univariate case considered throughout this work, the "target" state appearing in (6) can be identified with the m-th recentered ensemble analysis, U_T^m, written as

$$U_T^m = \alpha(U_A - \overline{U}_A) + U_A^m \,, \tag{22}$$

where, here, U_A corresponds to the hybrid analysis, U_A^m corresponds to the m-th EnSRF analysis, $\overline{U}_A$ corresponds to the EnSRF ensemble mean analysis, and the parameter $0 \leq \alpha \leq 1$ is introduced to allow for the possibility of "partial" recentering (e.g., Penny 2014). When $\alpha = 0$, no recentering takes place, the m targets become the member analyses themselves, and stability considerations fall under the traditional IAU context discussed earlier. When $\alpha > 0$, recentering takes a portion

of the hybrid analysis, and stability considerations fall under the Replay context examined earlier. Notice that here, part of the increment given to each ensemble member, $(U_A - \overline{U}_A)$, cannot be controlled by that member. It should also be simple to see that choosing $\alpha < 1$ has the effect of reducing the influence of the uncontrolled part of the increment, and thus help hold the instability in check. Incidentally, the Met Office has recently tried experiment with $\alpha < 1$, in a trial and error approach to try to avoid an instability of the kind discussed here, also plaguing their global system (G. Inverarity, pers. comm.).

Referring back to Fig. 14, once we realized the instability experienced in the parallel system was indeed due to how recentering is a form of Replay, we had the three choices discussed in Sec. 6 to try to avoid the instability. The most natural choice seemed to be that of applying a digital filter modulation to the IAU increments used to force the members of the ensemble. While doing so, recentering was also turned back on (with $\alpha = 1$). These changes are identified in Fig. 14 by the right-most (pink) vertical line, after which we see the surface pressure observation count increase slightly and completely stabilize. Indeed, after stabilization, there is an increase in the number of assimilated surface pressure observations as compared to the number of observations assimilated in the start of the time series shown in the figure; this is likely due to the improved accuracy of the corresponding background and analysis fields.

We should remark further that the instability developed in the hybrid system is truly independent of the flavor of hybrid the central DAS runs; the same behavior would have been observed had the 12.8 km system implemented a hybrid 3D-Var procedure relying on traditional (3D) IAU.

8. Implications to 4D assimilation procedures

A clear question to ask is whether the lessons learned here have implications in how 4DIAU-based procedures are implemented in practice. A detailed investigation with the objective to ad-

dress this question is planned for a follow up article. However, we close the present work by quickly answering the question in the affirmative and by providing an illustration supporting the claim.

The mechanism for using IAU in a 4D context is illustrated in Fig. 16, for a 6-hour assimilation window. The procedure is similar to that shown in Fig. 1(a) for 3D-Var except that now, the availability of *multiple* analyses from the underlying 4D variational procedure allows the Corrector step to make use of *multiple* incremental corrections to the model. At least two ways of incrementally correcting the model can be envisioned. The first, uses the increments from the underlying variational analysis, and the second uses the corresponding analyses calculated as updates to the corresponding background fields. The first procedure, referred to as "Fixed Increment" for simplicity, is more directly related with an incremental formulation of 4D approaches, the second, referred to as "Fixed Analysis", resembles a "nudging" approach where a tendency is created from each analysis within the assimilation time window by subtracting the analysis from the state of the model at a given time and keeping the corresponding tendency fixed over a certain time interval, typically one hour[2]. In both cases, the model is forced by piecewise-constant tendencies that change in time at a desired frequency. Lorenc et al. (2015) use the first approach combined with a nearest-time hourly implementation strategy. As pointed out The current GMAO FP system uses an hourly implementation of the, second, nudged-4DIAU approach.

Regardless of the approach taken, the presentation in Sec. 6c adds the possibility of applying a digital filter modulation to the IAU, or nudged-IAU, tendencies. An illustration of the error amplification factors for a few configurations of 4D strategies, and a 6-hour assimilation window, is given in Fig. 17. The current configuration of the GMAO FP system (blue curve) has an unstable mode peaking at about 90 minutes; furthermore, as indicated at the end of Sec. 5, this strategy has

[2]This is not to be confused with actual nudging, when the tendency is refreshed at each model time step.

the undesirable property of not tending to zero-error for the longest time scales. As a result, it does not take full advantage of the analysis and leaves model biases partially unchanged. The configuration corresponding to the implementation of Lorenc et al. (2015) (solid cyan curve) pushes the instability further into high frequencies and recovers the analysis in the long period range. As we have seen earlier, depending on the model effective damping time scales the instabilities seen in both these implementations might well be suppressed. The bottom panel of Fig. 9, indicates that the effective damping time scale of the C720 model used in FP is estimated at about 82 hours, thus the present GMAO FP hybrid configuration is not in danger of becoming unstable, but it would be safer and it could benefit from the better long-wave representation of the Lorenc et al. (2015) configuration. Even this latter can be improved upon when a digital filter is applied to modulate its tendencies. Applying a digital filter to the Lorenc et al. (2015) hourly configuration (dashed cyan curve) completely wipes out the instability but it has the likely adverse effect of not filtering as much as when no digital filter is applied. Indeed, it seems that a better configuration is obtained when the digital filter is used for a 3-hourly 4DIAU — that is, when the increments obtained at the two time edges and at the center of the assimilation window are used.

These are statements based purely on the filtering properties of the procedures. They do not inform us in any way which of these configuration leads to better fit to the observations and improved forecast skill scores. Though not the main goal of the present work, a glimpse of expected improvements to the current configuration of our FP system appears in the Figs. 18 and 19. Both present comparisons of the current nudged 4DIAU configuration of FP with the alternative configuration set to use a 3-hourly 4DIAU modulated with a digital filter. The experiments here are for a lower resolution configuration of the system than in FP: here, the hybrid analysis relies on a C360 ($\sim$ 25.6 km) forecast model, and on C90 ($\sim$ 100 km) ensemble members forecasts. Figure 18 compares the Northern Hemisphere Extratropics 500 hPa geopotential height anomaly corre-

lations (top) when using either one of the two configurations: FP-like (black curves); 3-hourly 4DIAU DF-modulated (red curves). For the sake of illustration, the result from a corresponding traditional IAU-based (non-hybrid) 3D-Var configuration is also shown (green curves). Difference of results from the 3-hourly 4DIAU-DF and 3D-Var with those from the FP-like configuration are shown in the bottom panel, with associated significance bars. It is quite apparent that the current FP-like configuration is an improvement over traditional IAU-based 3D-Var, but it is also clear that there is an equally significant improvement to be made by reconfiguring the Corrector step of the DAS using what we learned in the present work.

The results of Fig. 18 are from self verification and can always be misleading. Figure 19 provides a relative comparison, of selective scores, when forecasts from the FP-like and 3-hourly 4DIAU-DF experiments are verified against NCEP analyses (top) and observations (bottom). The relative scores are calculated such that negative values reflect improvements obtained with the 3-hourly 4DIAU-DF configuration as compared to the FP-like settings. The thin bars are 90% statistical significance. Against both observations and NCEP analyses, the reconfiguration of IAU in the 4D system leads to an overall improvement with respect to the current FP-like settings.

9. Conclusions

The present work takes a second look at the Incremental Analysis Update (IAU) procedure of Bloom et al. (1996) after identifying previously unnoticed instabilities in the procedure. Our evaluation reveals that oscillations with periods longer than the analysis interval are stable, but those of shorter period may actually grow. In particular, those with periods shorter than the analysis interval but longer than half the assimilation interval are found to be unstable, with fastest growth occurring for a period just longer than two-thirds of the interval.

Two applications of IAU are considered in this work. One is when IAU is implemented in a cycling data assimilation system (DAS); another is when IAU is used to, for example, downscale a given set of analyses with a model running at higher resolution than the one used to produce the original analyses. The latter mode of IAU is referred to in this work as "Replay" and its applications are broader than simply downscaling exercises, as it allows for testing model changes without having to run a complex DAS.

In the context of Replay, we show that the effective damping of general circulation models plays a fundamental role in controlling the development of the instability. At low resolutions the model is rather dissipative and is able to damp the IAU instability. As resolution increases, and the effective model dissipation decreases, the risk of exciting the instability increases. Indeed, GMAO first stumbled on the instability when trying to downscaling the MERRA-2 analyses with a 7 km model. A temporary solution was found then to allow the downscaling to proceed.

The work here provides a broad understanding of the IAU instability and finds at least three approaches to avoid it. A so-called "Sweet-Spot" approach examines an IAU stability diagram and finds an "ideal" configuration for the Predictor and Corrector steps of IAU. A second approach (taken in the downscaling project), proposes to perform a "Background Averaging" at given frequency in order to calculate IAU tendencies. And finally a third approach, proposes to modulate the IAU tendencies with weights from a digital filter. This latter, is the simplest and most general solution to avoid the instability. Indeed, this turns out to be a nice application of the work of Polavarapu et al. (2004) on the relationship between IAU and digital filter.

The IAU instability can arise in both Replay and regular DAS settings. In DAS settings we find the instability to have slightly smaller amplification factors, for given model dissipation rates. This explains why DAS applications such as MERRA-2, at roughly 50 km, are stable. However if, without knowledge of the findings here, we were to increase the model resolution of the current

GMAO forward processing (near-real time) DAS, from its 12.6 km to roughly 7 km we would likely trigger the instability, just as it was triggered in the first attempt to downscale MERRA-2 to 7 km.

But more interesting is the relationship between Replay and recentering of the members of the ensemble around the hybrid analysis in a hybrid DAS scheme. We explain here how recentering is a type of Replay in the sense that part of the increment used to create the IAU tendency is not controlled by the ensemble analysis (but rather *only* by the hybrid analysis). We provide an incidental example of how an IAU-based ensemble DAS used to feed into a hybrid DAS can develop the instability. We also show the instability can be completely eliminated by modulating the IAU tendencies driving the ensemble member forecasts with a digital filter.

Finally, the present work provides a brief insight on what the findings here mean for configuring 4D hybrid systems using 4DIAU. Although, a full study on this matter is postponed to a follow up article, we provide a brief example of the potential for possible improvements in the GMAO hybrid 4D-EnVar system by simply reconfiguring its current 4DIAU settings along the lines of the findings in this work.

Acknowledgements. This work was supported under the GMAO Core funding of NASA's Modeling, Analysis and Prediction (MAP) program. The computational work was carried out on the Linux Discover System through cooperation with the NASA Center for Climate Simulation at Goddard Space Flight Center. Thanks to Saroja Polavarapu and two other anonymous reviewers for providing insightful comments and help us improve the readability of the manuscript.

References

Balmaseda, M. A., A. Vidard, and D. L. T. Anderson, 2008: The ECMWF ocean analysis system: ORA-S3. *Mon. Wea. Rev.*, **136**, 3018–3034.

Bloom, S. C., L. L. Takacs, and E. Brin, 1991: A scheme to incorporate analysis increments gradually in the GLA assimilation system. *Ninth Conf. on Numerical Weather Prediction*, 110–112.

Bloom, S. C., L. L. Takacs, A. M. da Silva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Mon. Wea. Rev.*, **124**, 1256–1271.

Buehner, M., et al., 2015: Implementation of deterministic weather forecasting systems based on ensemble-variational data assimilation at Environment Canada. Part I: The global system. *Mon. Wea. Rev.*, **143**, 2532–2559.

Carton, J. A. and B. S. Giese, 2008: A reanalysis of ocean climate using simple ocean data assimilation (SODA). *Mon. Wea. Rev.*, **136**, 2999–3017.

Gelaro, R. and coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *J. Climate*, **30**, 5419–5454.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability*. 1st ed., Cambridge University Press, Cambridge, U.K.

Kleist, D. T., 2012: An evaluation of hybrid variational-ensemble data assimilation for the NCEP GFS. Ph. d. thesis, University of Maryland, 163 pp pp. URL `http://www.emc.ncep.noaa.gov/gmb/wd20dk/docs/phd/DarylKleist\_PhDThesis\_Revised.pdf`.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. Lord, 2009: Introduction of the GSI into the NCEP Global Data Assimilation System. *Wea. Forecasting*, **24**, 1691–1705, doi:http://dx.doi.org/10.1175/2009WAF2222201.1.

Lawless, A., 2010: A note on the analysis error associated with 3D-FGAT. *Q. J. Royal Meteorol. Soc.*, **136(B)**, 1094–1098.

Lee, M.-S., Y.-H. Kuo, D. M. Barker, and E. Lim, 2006: Incremental analysis updates initialization technique applied to 10-km MM5 and MM5 3DVAR. *Mon. Wea. Rev.*, **134**, 1389–1404.

Lei, L. and J. S. Whitaker, 2016: A four-dimensional Incremental Analysis Update for the ensemble Kalman filter. *Mon. Wea. Rev.*, **144**, 2605–2621.

Lorenc, A. C., N. E. Bowler, A. M. Clayton, and S. R. Pring, 2015: Comparison of Hybrid-4DEnVar and Hybrid-4DVar Data Assimilation Methods for Global NWP. *Mon. Wea. Rev.*, **143**, 212–229.

Lorenc, A. C., et al., 2000: The Met. Office global three-dimensional variational data assimilation scheme. *Q. J. Royal Meteorol. Soc.*, **126(B)**, 2991–3012.

Lynch, P. and X.-Y. Huang, 1992: Initialization of the HIRLAM model using a digital filter. *Mon. Wea. Rev.*, **120 (6)**, 1019–1034, doi:10.1175/1520-0493(1992)120⟨1019:IOTHMU⟩2.0.CO;2.

Massart, S. M., B. Pajot, A. Piacentini, and O. Pannekoucke, 2010: On the merits of using a 3D-FGAT assimilation scheme with an outer loop for atmospheric situations governed by transport. *Mon. Wea. Rev.*, **138**, 4509–4522.

Orbe, C., L. D. Oman, S. E. Strahan, D. W. Waugh, S. Pawson, L. L. Takacs, and A. M. Molod, 2017: Largescale atmospheric transport in GEOS replay simulations. *J. Advances Modeling Earth Systems*, **9**, 2545–2560, doi:10.1002/2017MS001053.

Ourmières, Y., J.-M. Brankart, L. Berline, P. Brasseur, and J. Verron, 2006: Incremental analysis update implementation into a sequential ocean data assimilation system. *J. Atmos. Oceanic Tech.*, **23**, 1729–1744.

Penny, S. G., 2014: The hybrid local ensemble transform Kalman filter. *Mon. Wea. Rev.*, **142**, 2139–2149.

Polavarapu, S., S. Ren, A. Clayton, D. Sankey, and Y. Rochon, 2004: On the relationship between incremental analysis updating and incremental digital filtering. *Mon. Wea. Rev.*, **132**, 2495–2502.

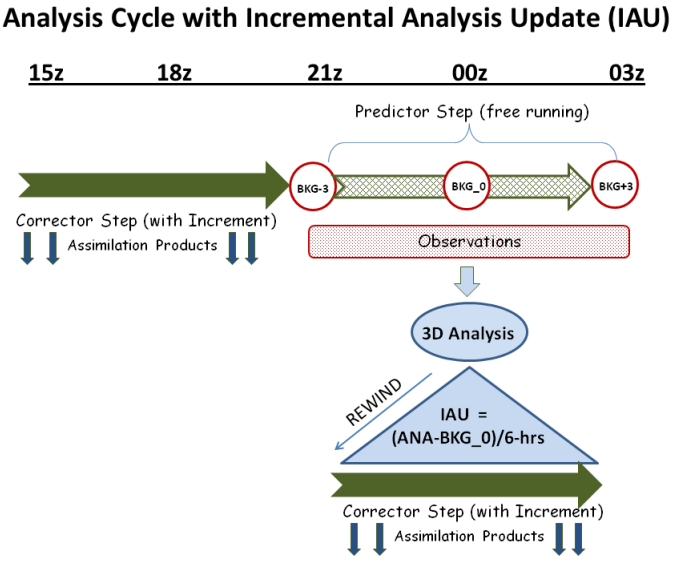
Schubert, S. D., R. D. Rood, and J. Pfaendtner, 1993: An assimilated dataset for earth science applications. *Bull. Amer. Meteorol. Soc.*, **74**, 2331–2342.

Todling, R. and A. El Akkraoui, 2018: The GMAO Hybrid Ensemble-Variational Atmospheric Data Assimilation System: Version 2.0. NASA Tech. Memo. 104606-50, NASA Goddard Space Flight Center.

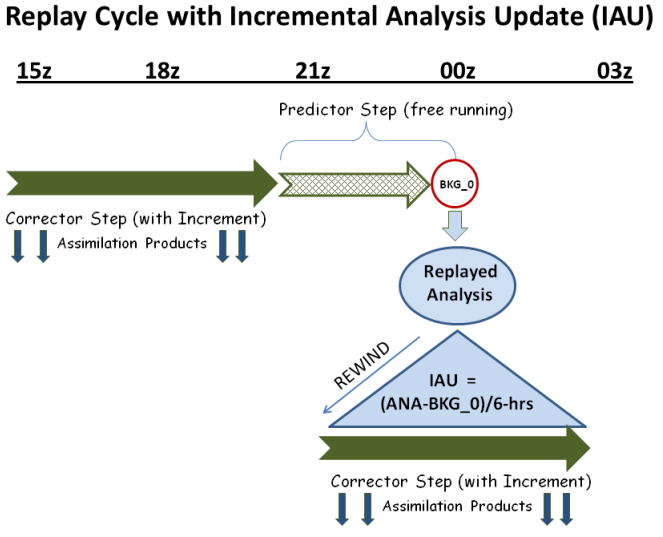
Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP Global Forecast System. *Mon. Wea. Rev.*, **136**, 463–482.

LIST OF FIGURES

(a)



(b)



693 FIG. 1. Schematic representation of the Incremental Analysis Update (IAU) as implemented in: (a) a fully-
694 cycled 3D-FGAT DAS, such as MERRA-2; and (b) a Replay procedure to force AGCM experiments to track
695 pre-existing analyses, such as those from MERRA-2.

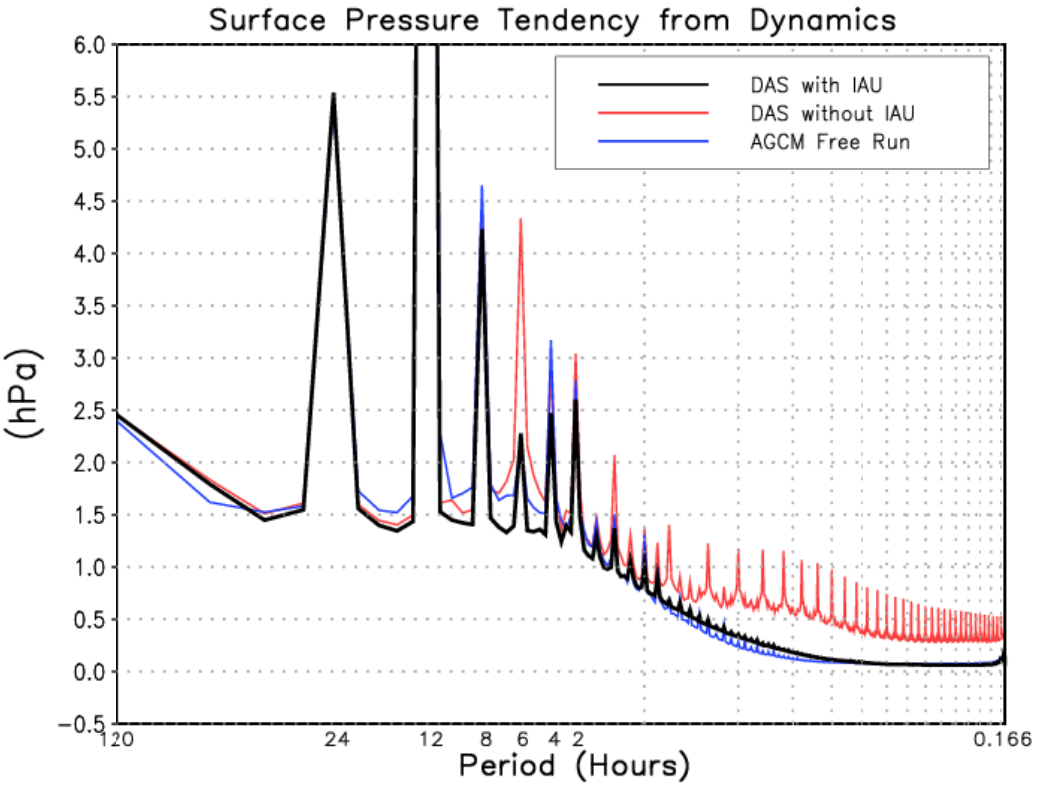
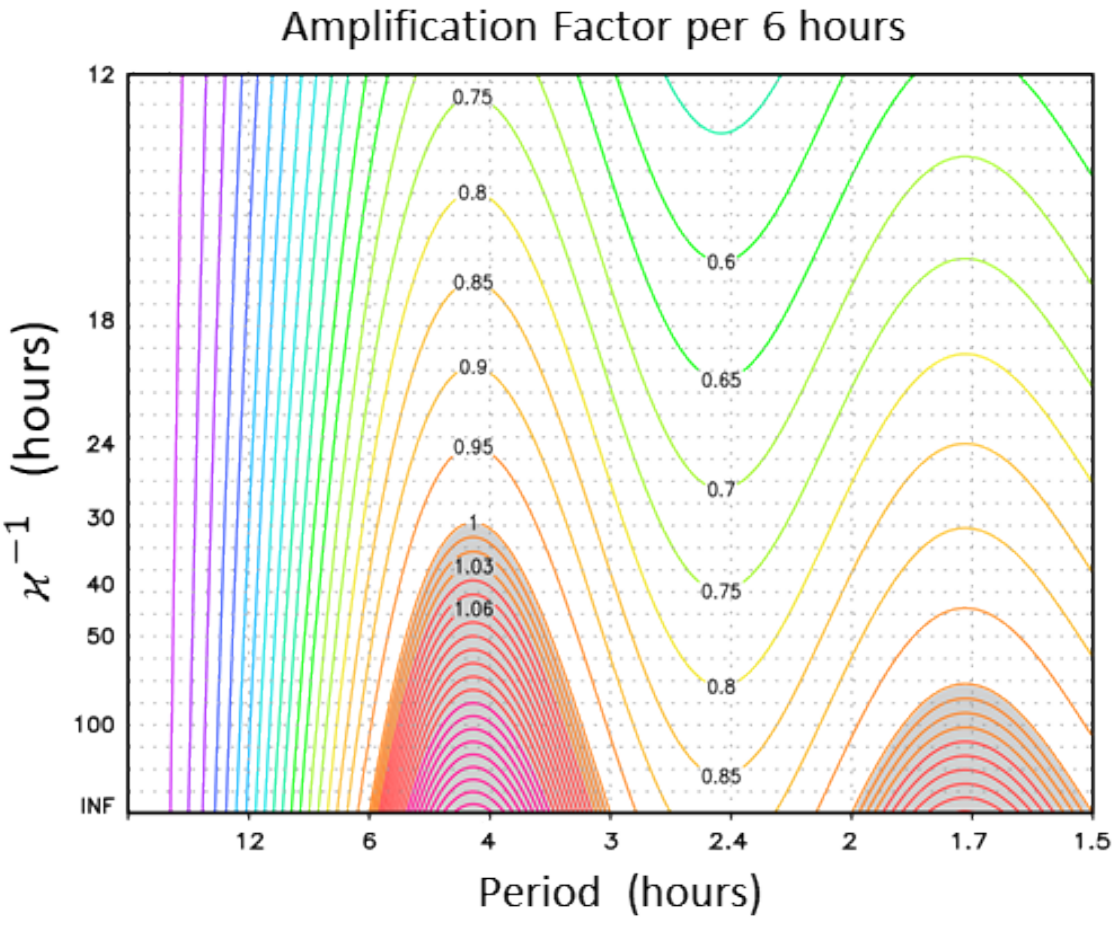


FIG. 2. Comparison of the tropically-averaged (30S-30N) power spectrum amplitudes of the surface pressure tendency (hPa/day) due to dynamics between a DAS with direct insertion, a DAS using IAU, and an AGCM free run.



699 FIG. 3. Amplification factor for the Replay case ($K = 1$) as a function of the period and damping time scale,
700 κ^{-1}, of the oscillation. This is for the standard case of a 3-hour Predictor and a 6-hour Corrector. Stable regions
701 (values ≤ 1) are contoured using 0.05 intervals, while unstable regions (shaded, with values > 1) are contoured
702 using 0.01 intervals.

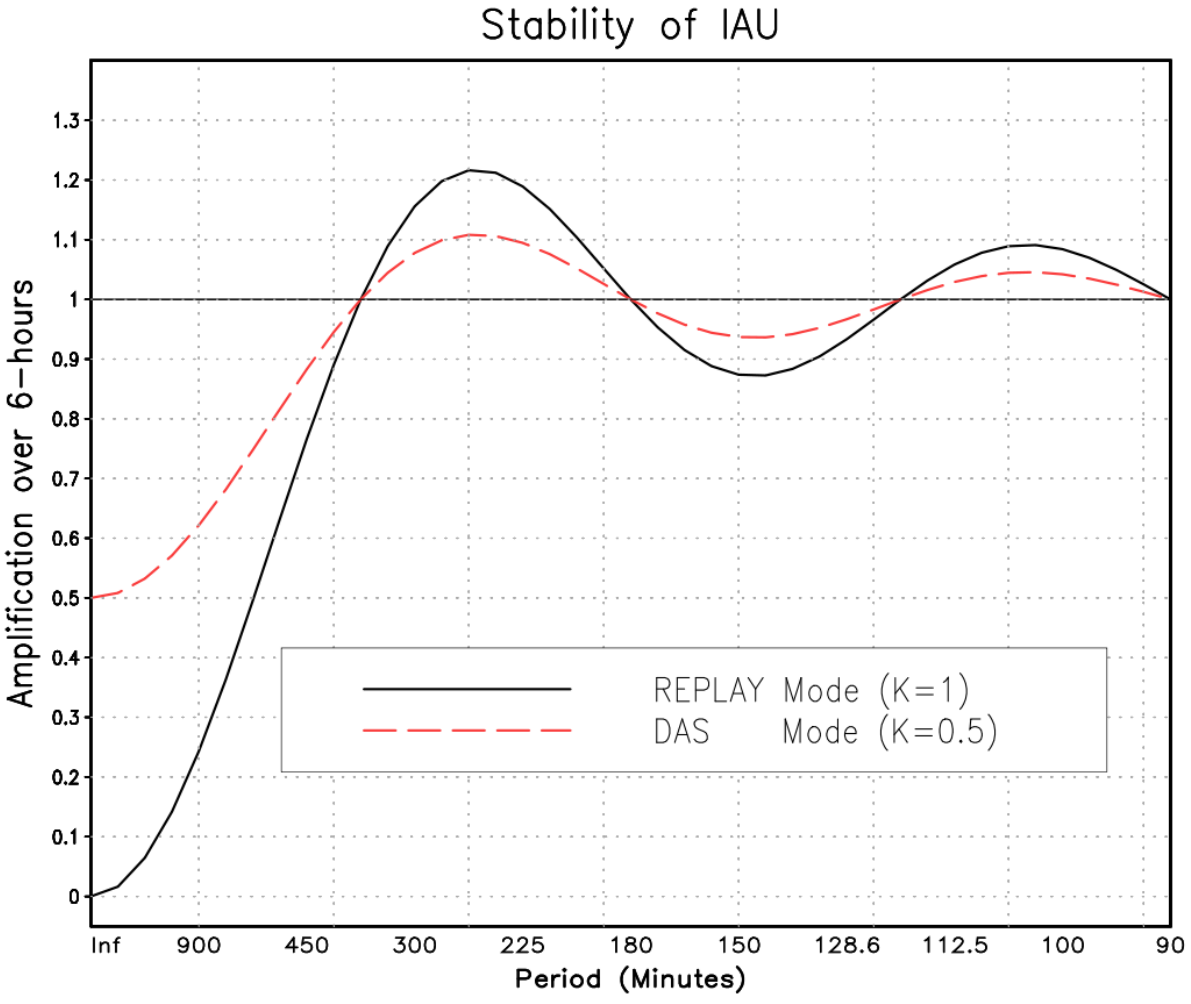


FIG. 4. Amplification factors as a function of period. In our simple model, Replay mode corresponds to $K = 1$. The IAU-based DAS mode corresponds to $K < 1$. The damping effect of the DAS is illustrated for $K = 0.5$. Both curves are for non-dissipative dynamics, $\kappa = 0$.

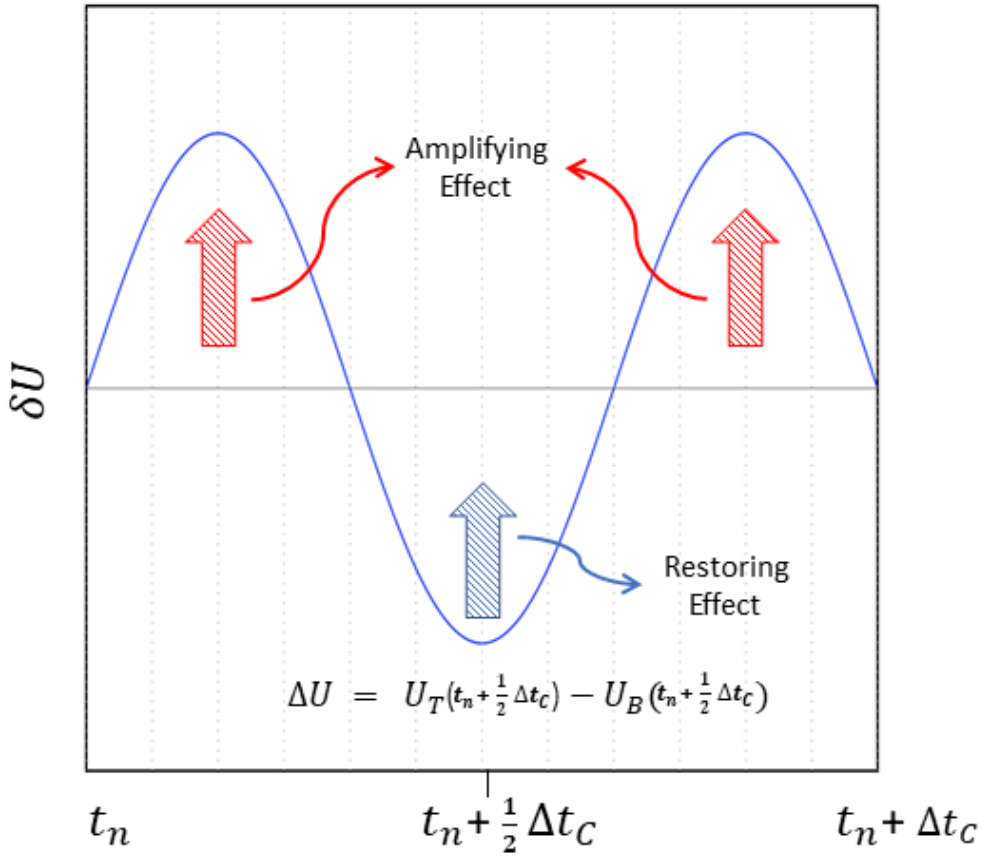
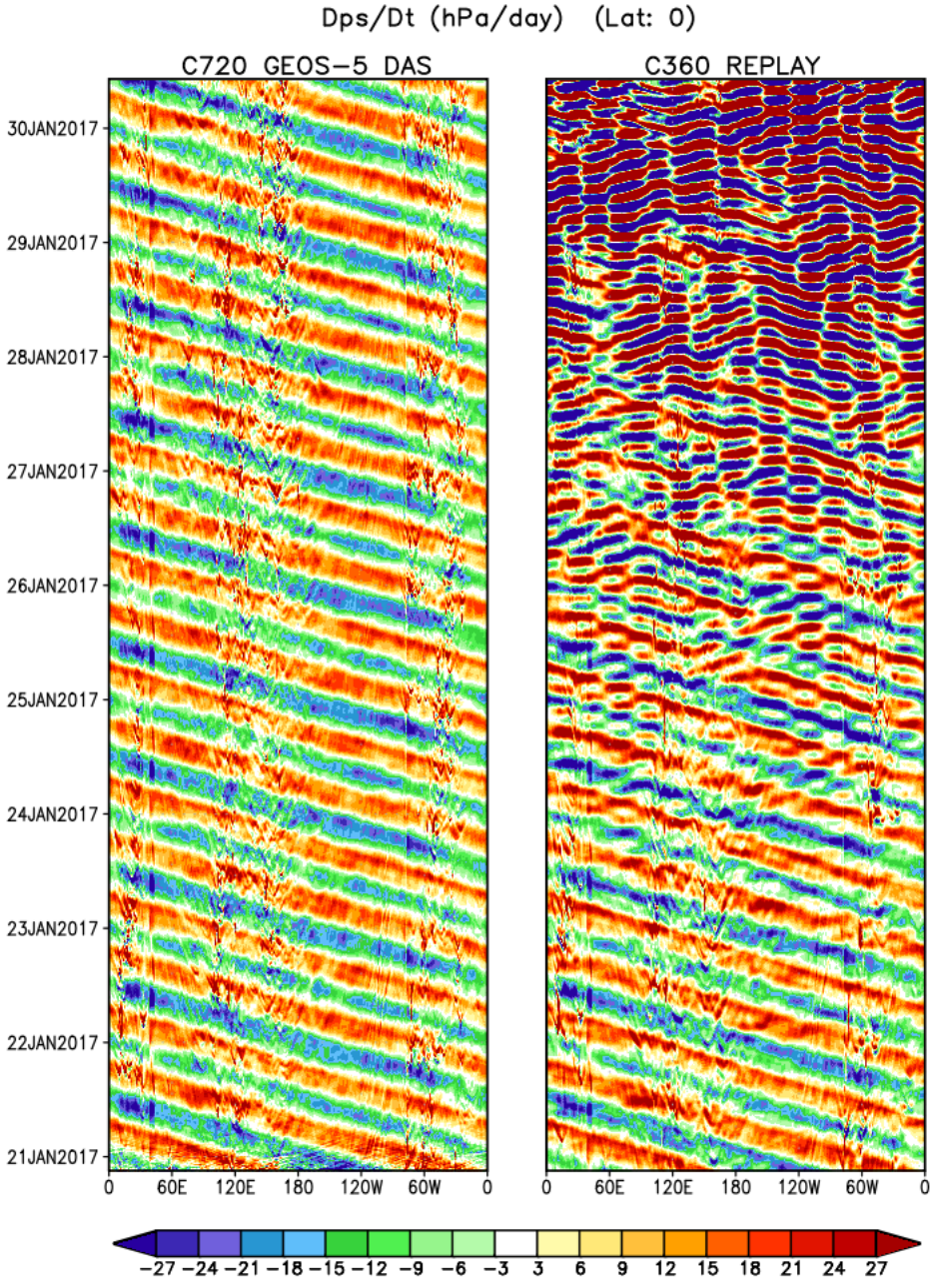
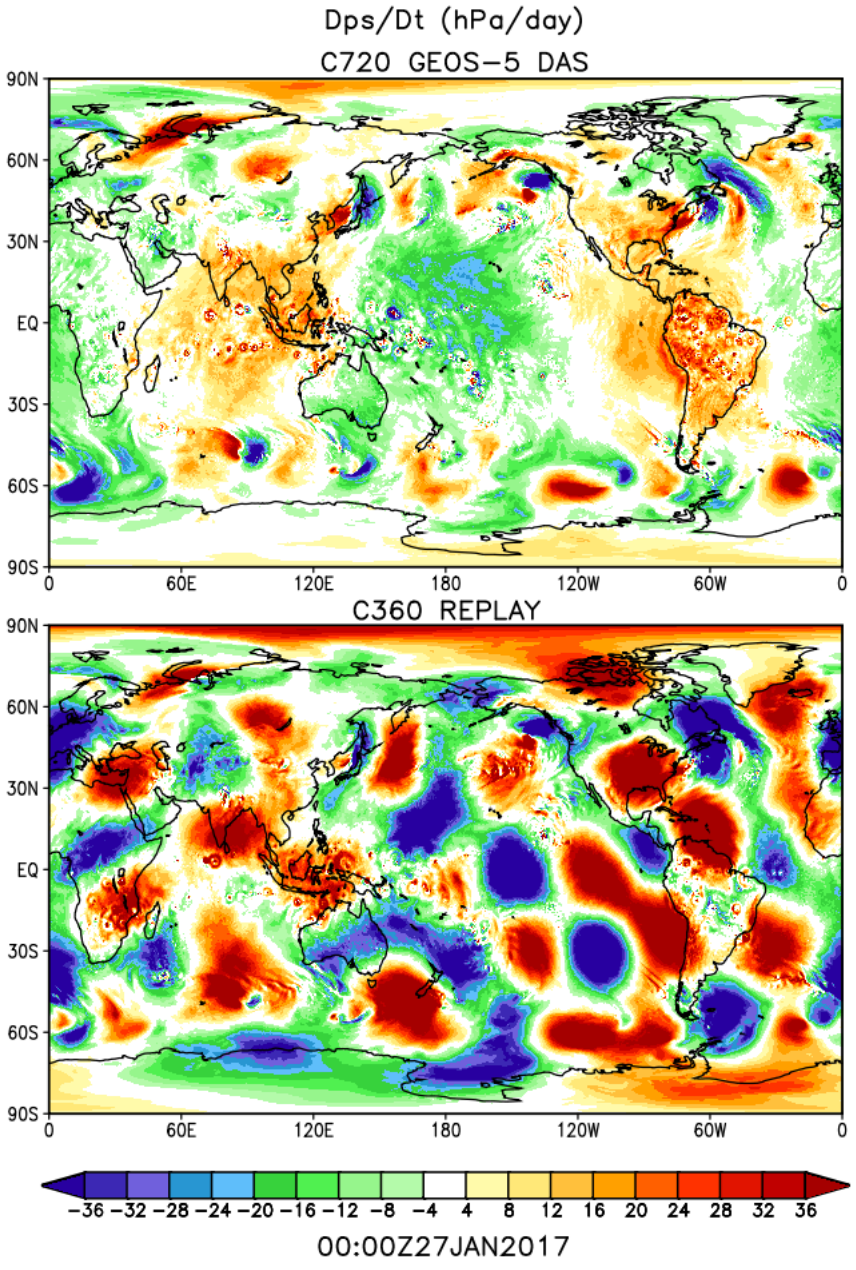


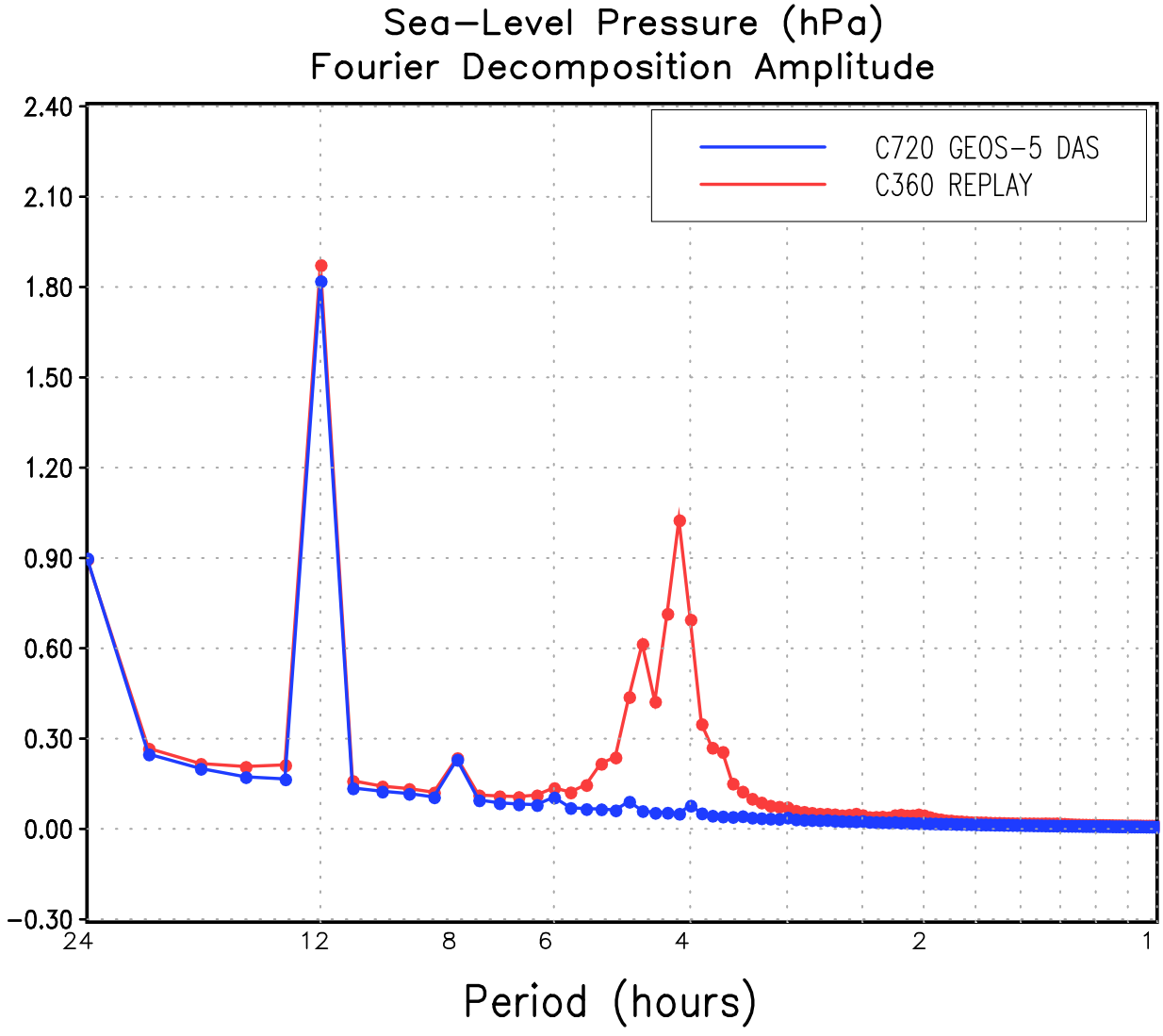
FIG. 5. Schematic of the instability of IAU when the increment is based on the midpoint of the analysis cycle. Shown is the departure of the free solution from the target solution, δU, for an oscillation with a period of $\frac{2}{3}\Delta t_C$. The sense of the increment computed at the mid-point, ΔU, is shown by the arrows. When this increment is applied uniformly over the entire interval, it will only be restorative (driving δU to zero) over the middle third of the interval.



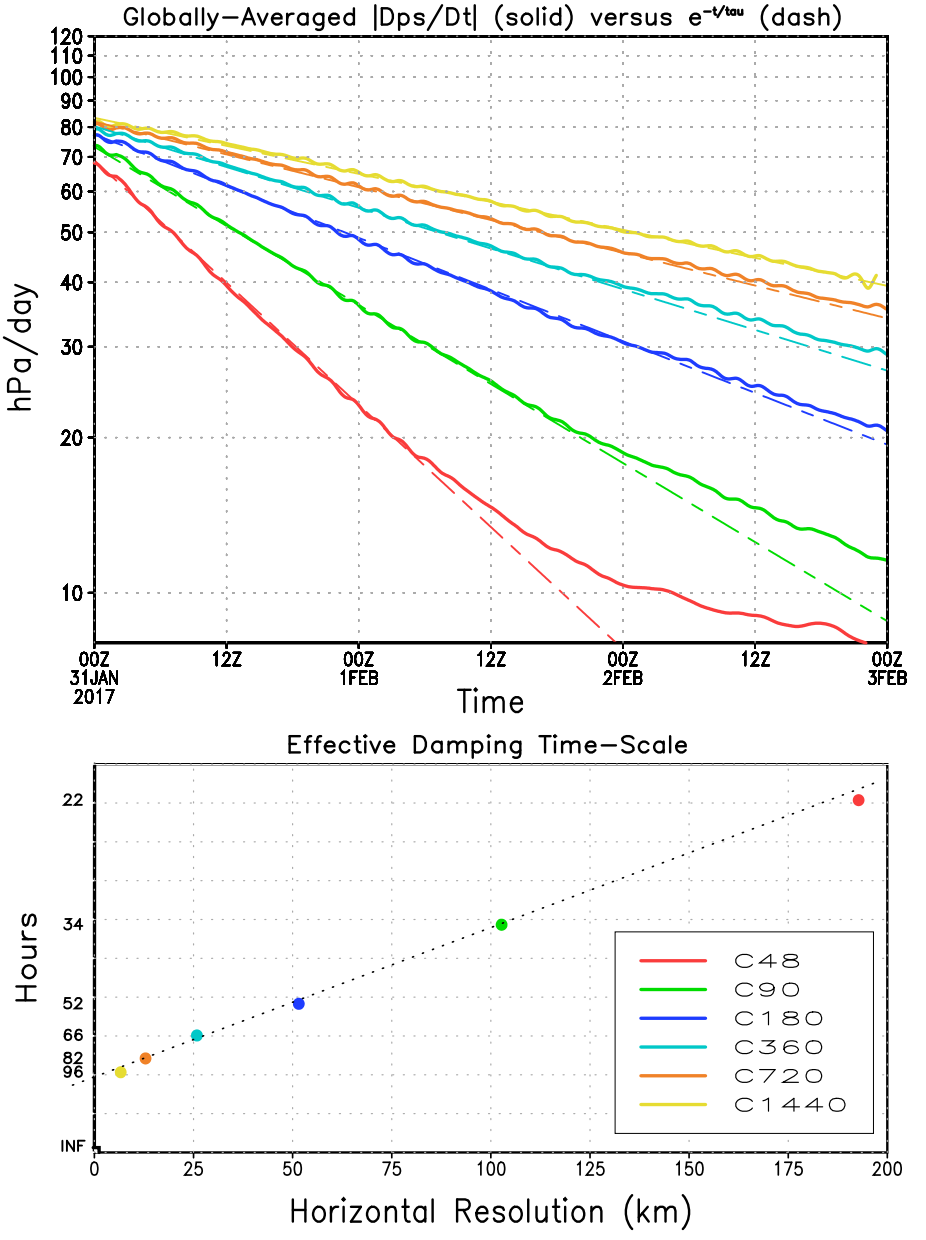
711 FIG. 6. Time evolution of the surface pressure tendency due to dynamics (hPa/day) at the equator from the
712 C720 GMAO DAS (Cubed grid with a nominal resolution of $\sim$ 12.8 km, left panel) and the standard C360
713 REPLAY at half the resolution (right panel).



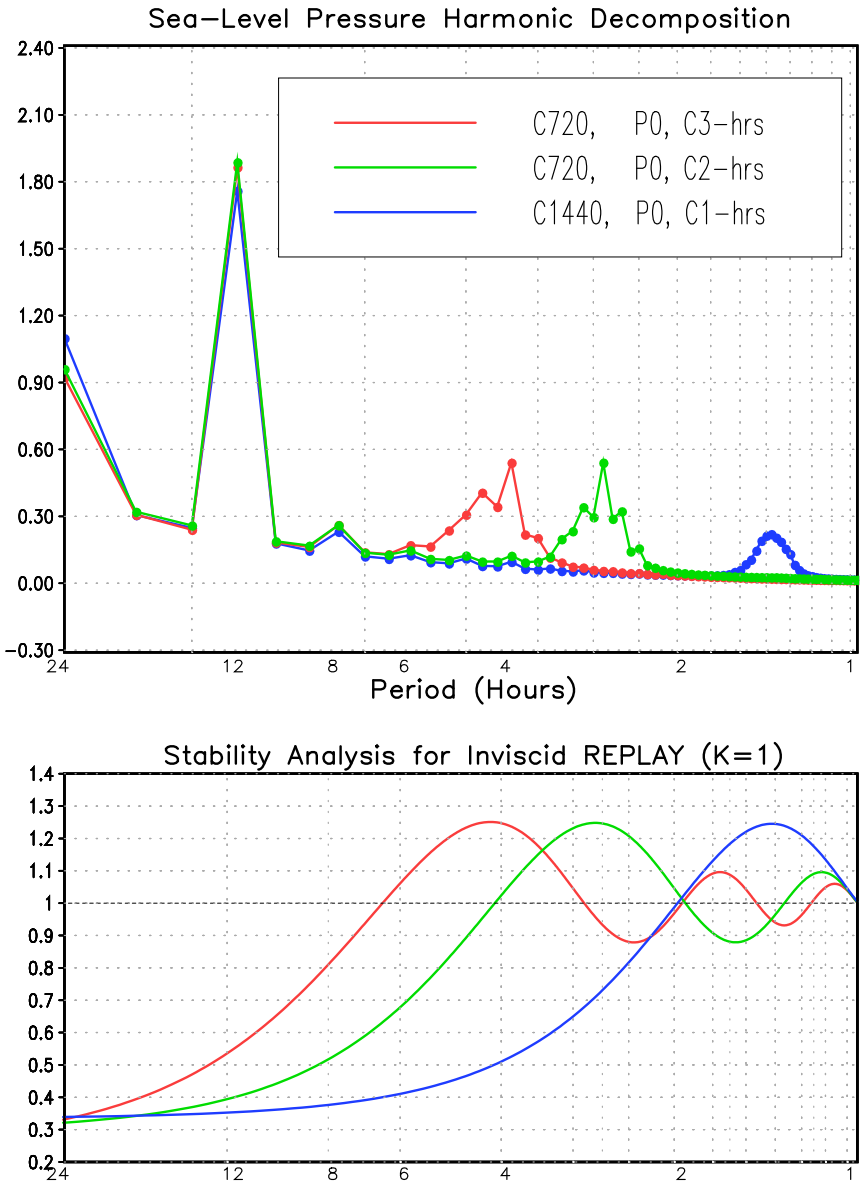
714 FIG. 7. Instantaneous snapshots of the surface pressure tendency due to dynamics (hPa/day) on January 27,
715 2017 from the C720 DAS and the standard C360 Replay.



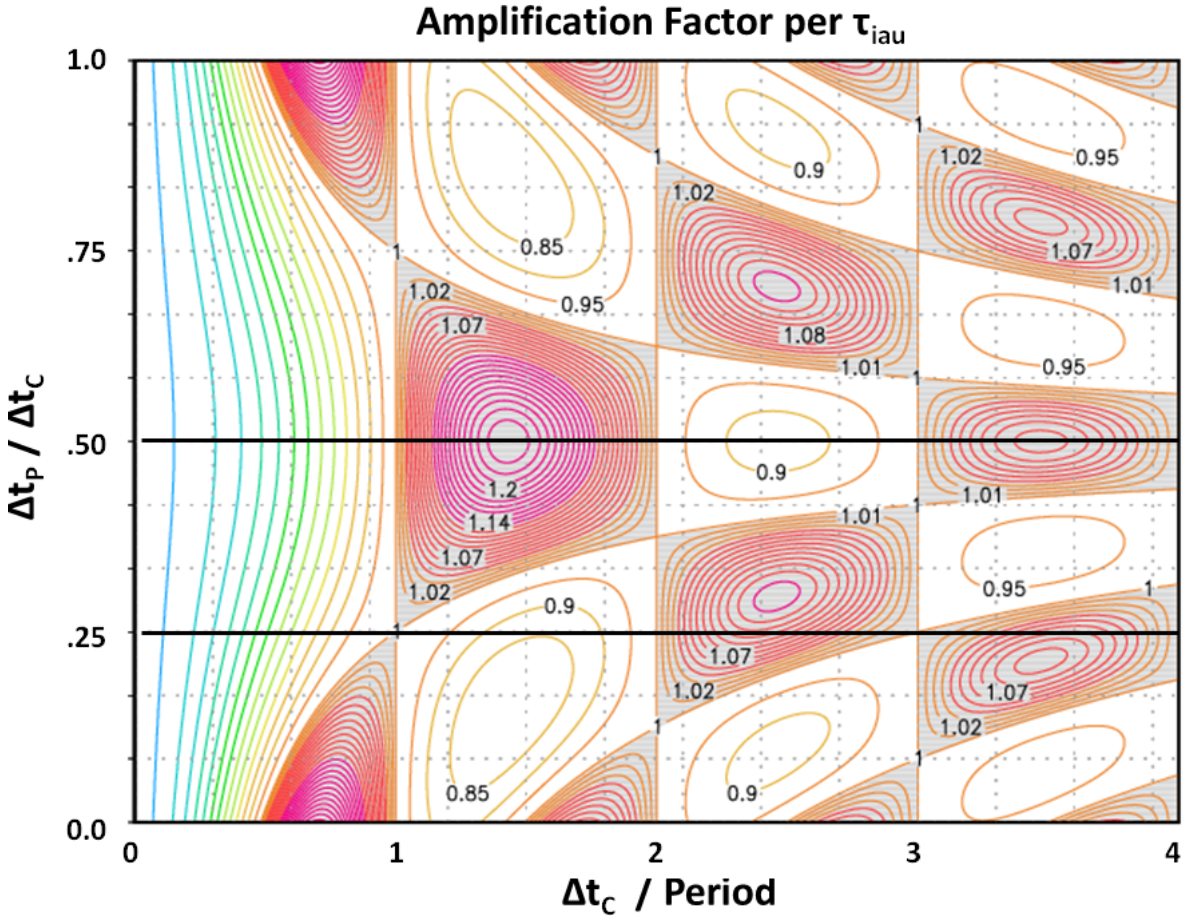
716 FIG. 8. Tropically-averaged (30S to 30N) Fourier decomposition of sea-level pressure computed from January
717 25-29 data with an output frequency of 5-minutes.



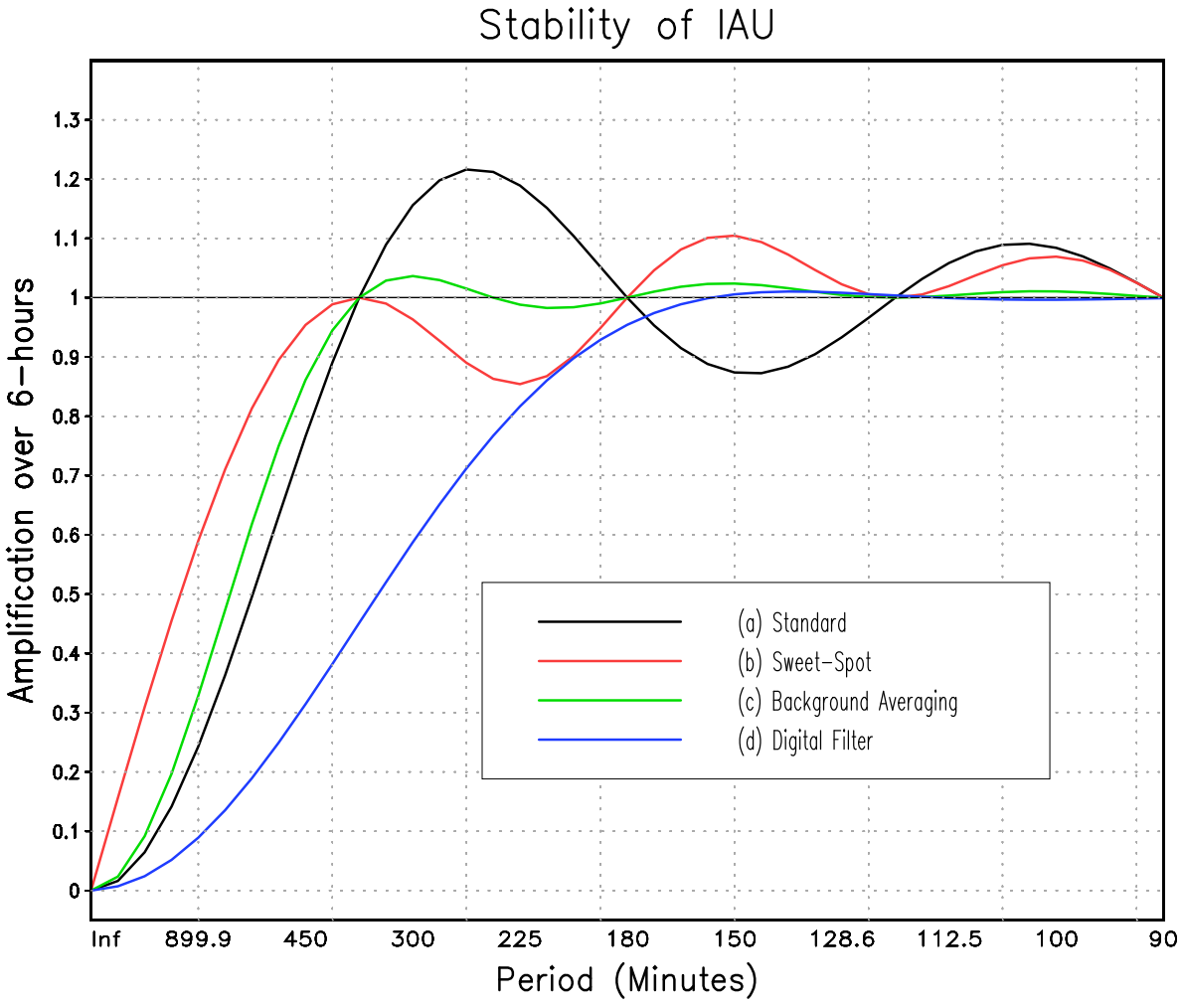
718 FIG. 9. Upper panel: Globally-averaged magnitude of the surface pressure tendency compared with e-folding
719 curves. Lower panel: Asymptotic behavior of effective GEOS AGCM damping time-scale of the unstable 4-hour
720 Replay mode as a function of horizontal resolution.



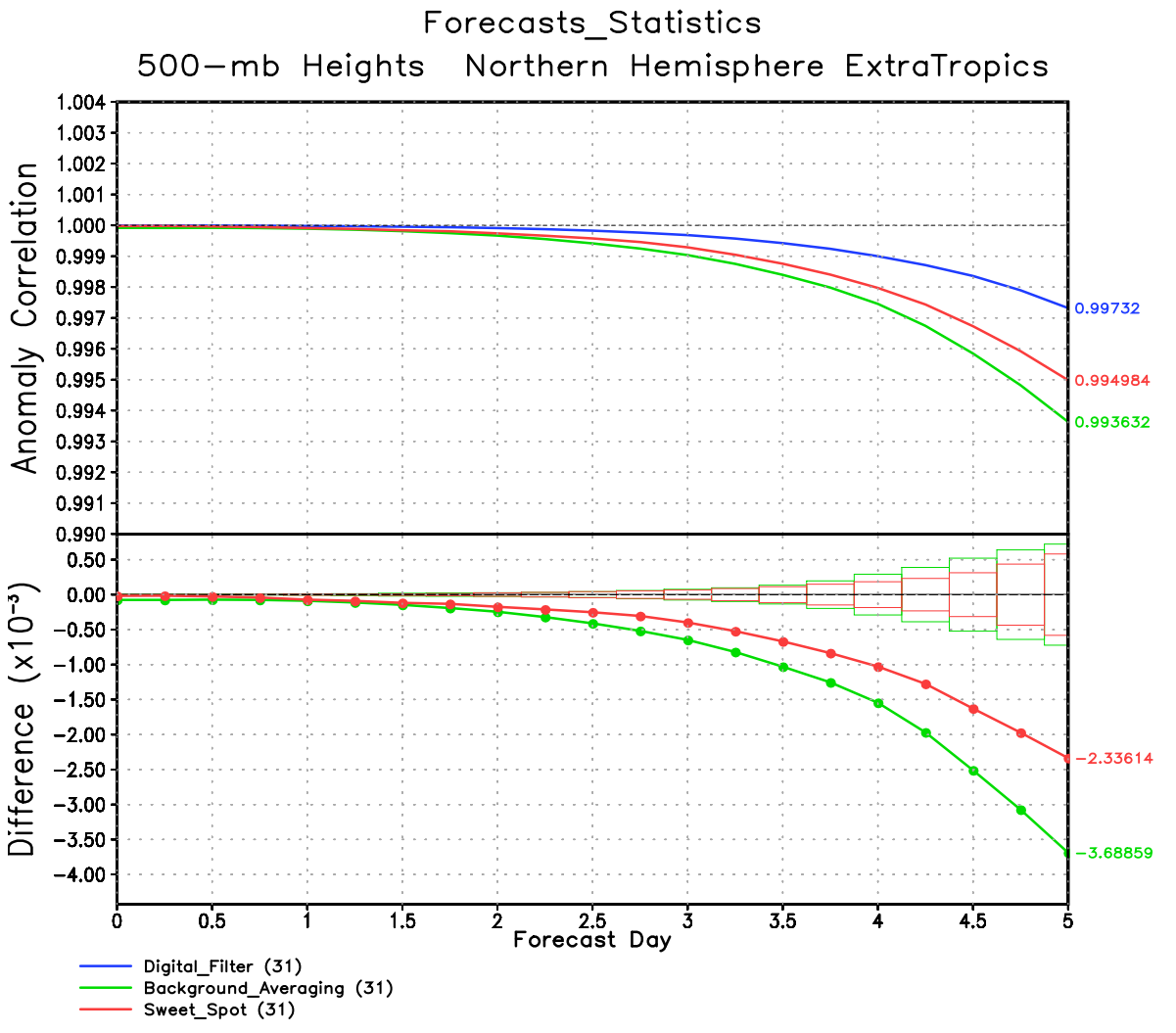
721 FIG. 10. Sea-level pressure harmonic decomposition from C720 and C1440 Replays using various Predictor
722 and Corrector durations. The lower panel depicts the stability analysis of the error associated with the inviscid,
723 Replay $(K = 1)$ case based on (10).



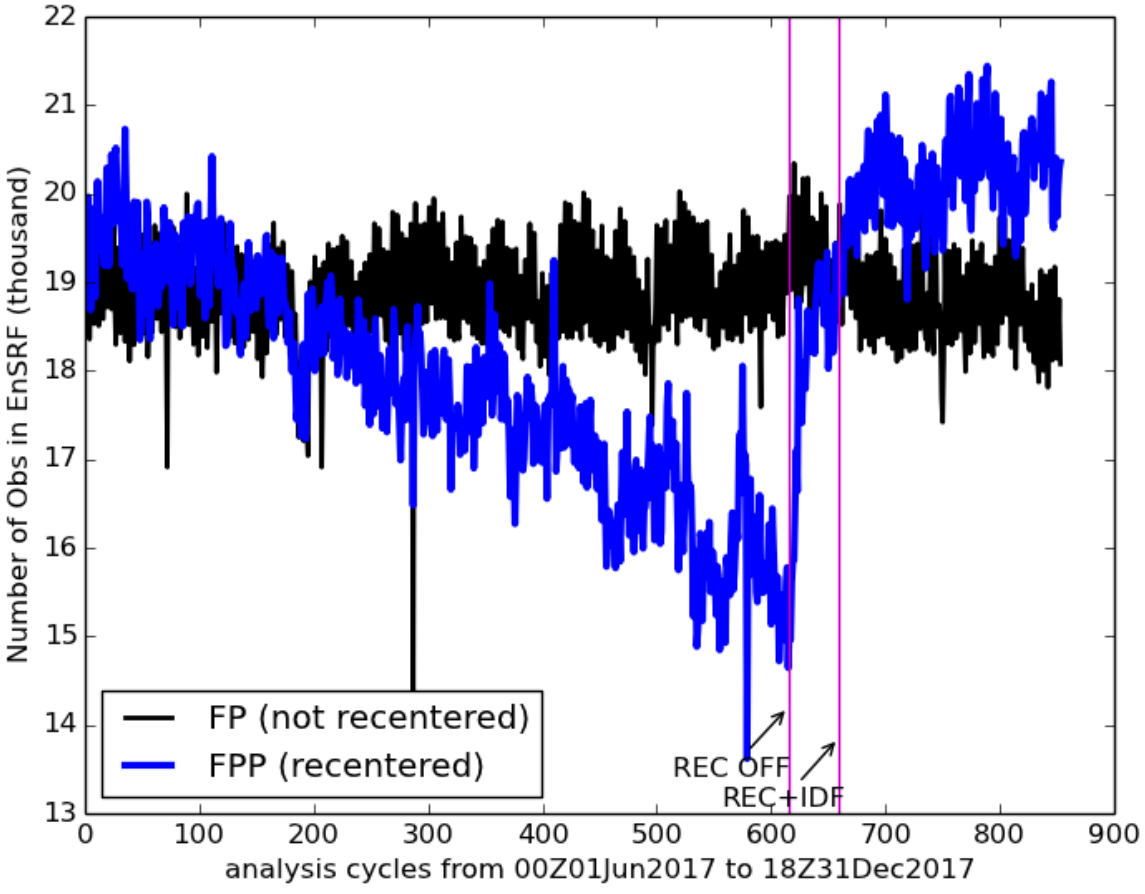
724 FIG. 11. The Amplification factor for a Replay $(K = 1)$ of length τ_{iau}, as a function of normalized frequency
725 and the ratio of the Predictor to Corrector duration. The plot assumes $\kappa = 0$.



726 FIG. 12. The Amplification factor as a function of period for (a) the Standard 3-hr Predictor 6-hr Corrector
727 IAU strategy, (b) the 1.5-hr Predictor 6-hr Corrector "Sweet-Spot" strategy, (c) the Standard IAU strategy with
728 Background Averaging of 4-hours, and (d) the Standard IAU strategy with the Digital Filter.



729 FIG. 13. The Northern Hemisphere Extratropics 500-mb Height Anomaly Correlations from 5-day forecasts
730 using the three methods for controlling the IAU/Replay instability.



731 FIG. 14. Count of surface pressure observations assimilated in the EnSRF analysis of the Hybrid 4D-EnVar
732 system over the course of six months. The black curve is for the non-recentered forward-processing (FP) system;
733 the blue curve is for the parallel (FPP) system, in which recentering was being reinstated. The vertical pink lines
734 indicate the instants when recentering was turned off (REC-OFF); and when recentering was turned back on
735 together with a digitial filter modulation applied to the 3DIAU ensemble integration (REC-IDF).

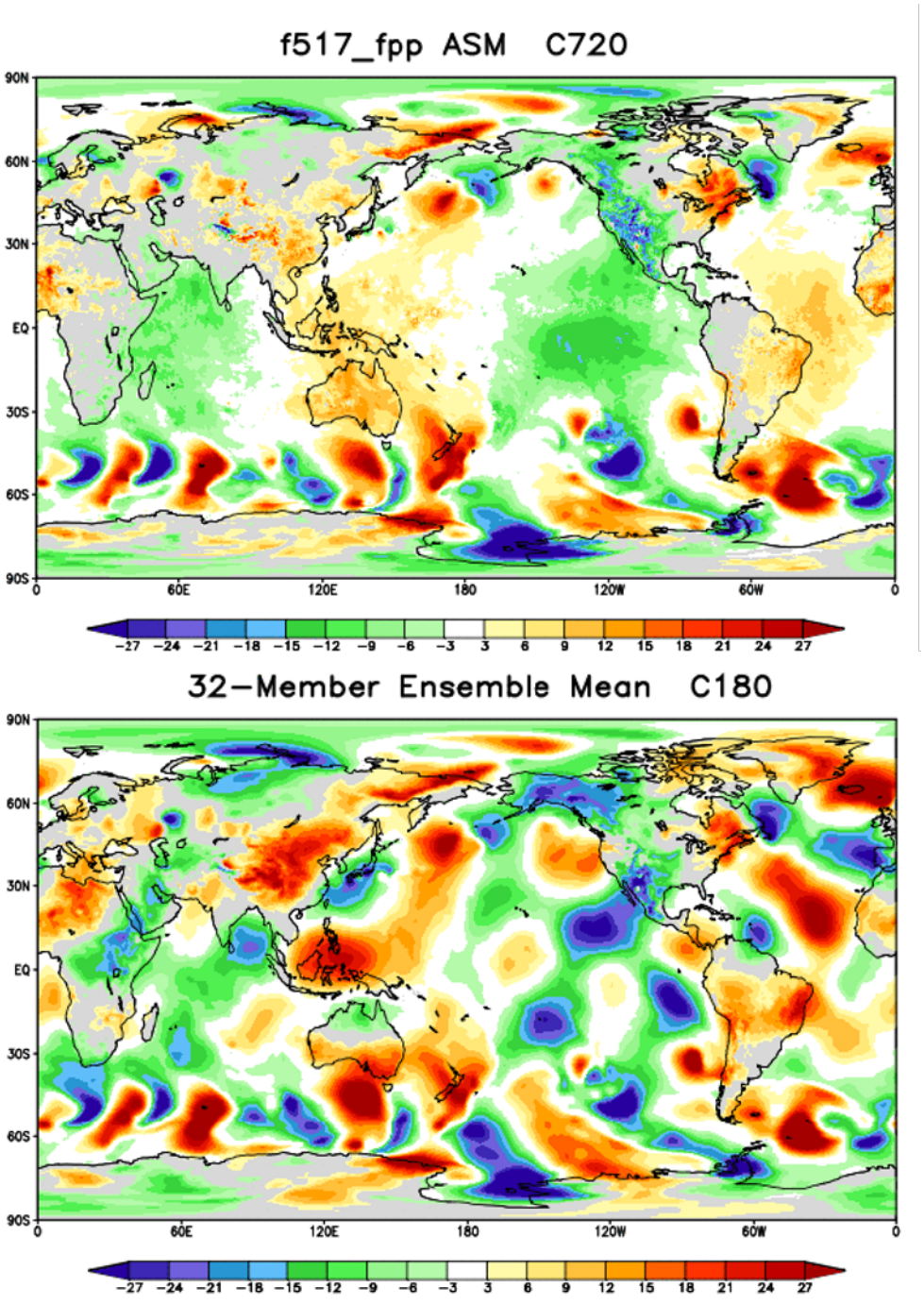
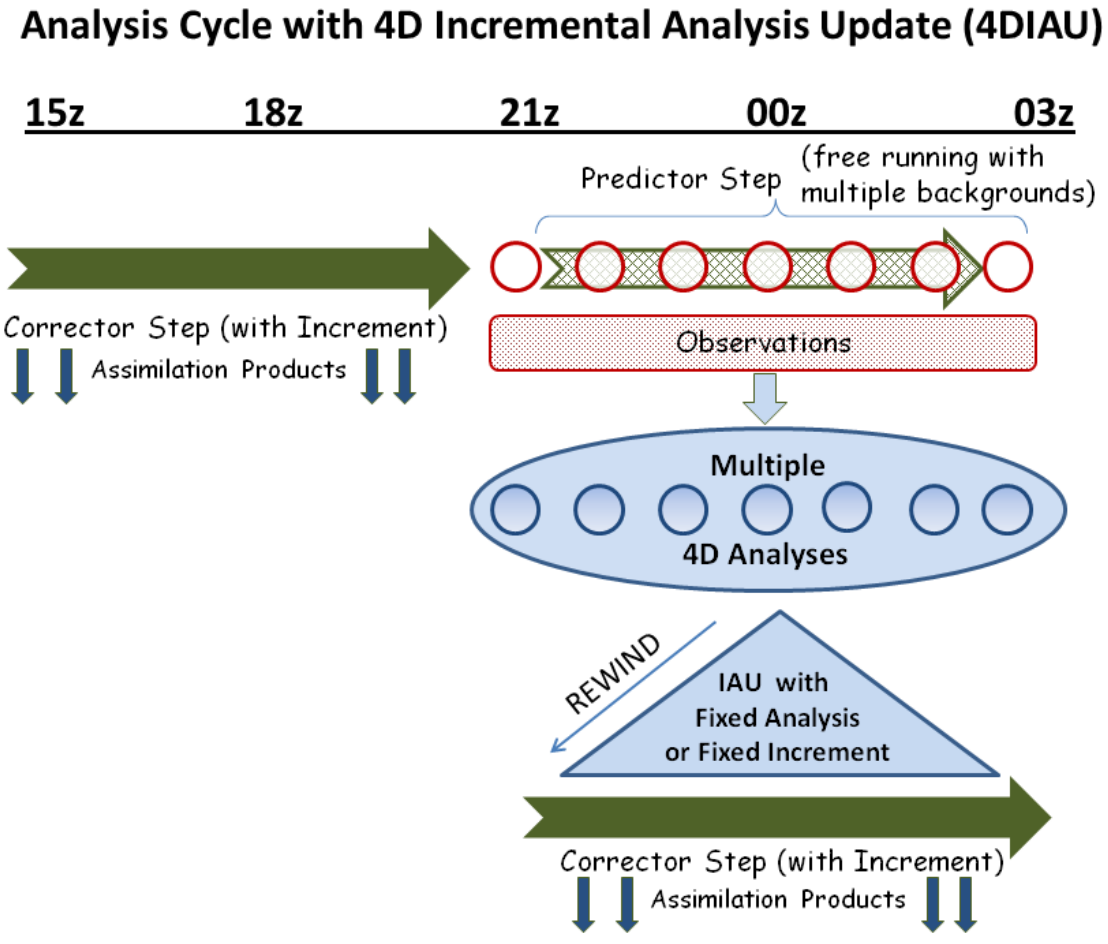
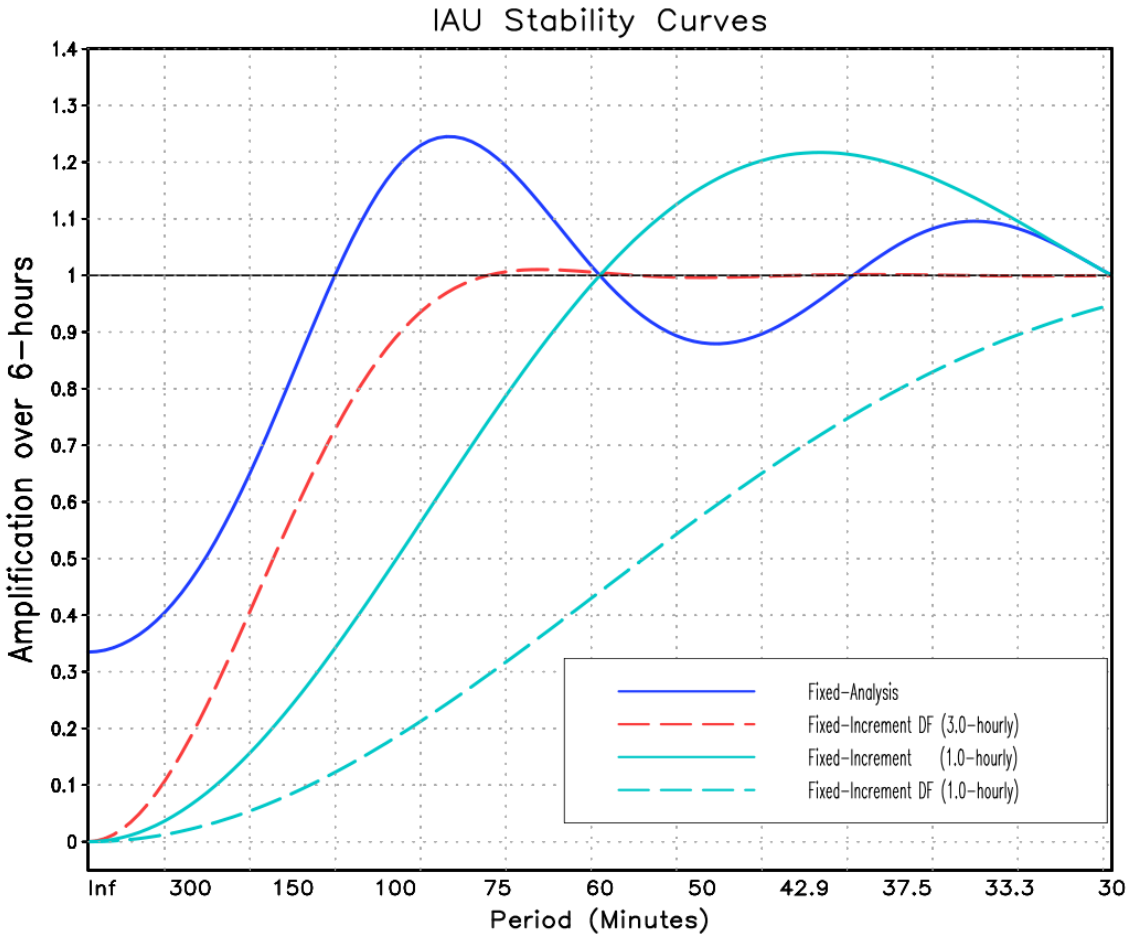


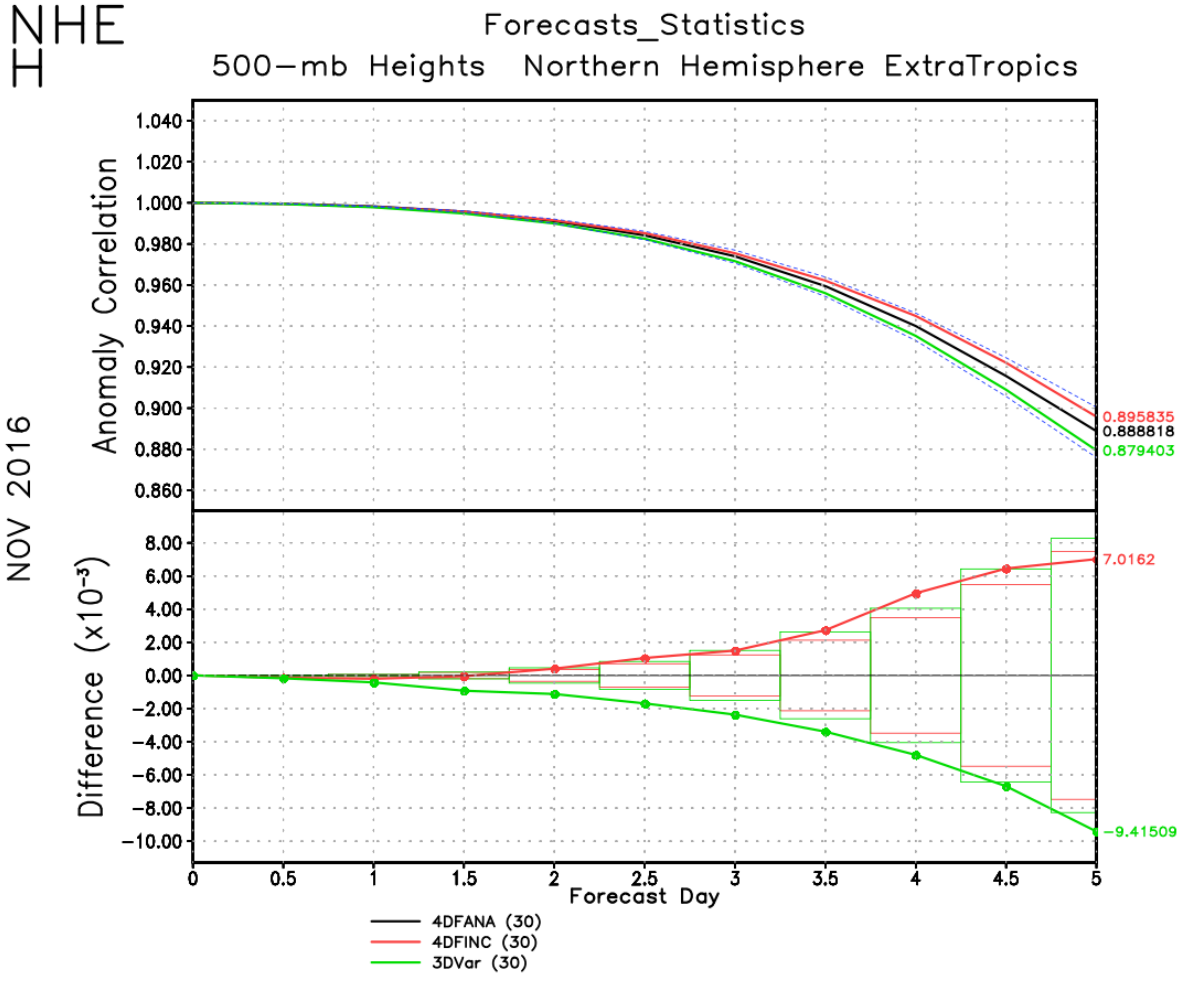
FIG. 15. Manifestation of the IAU instability in the ensemble members of the GMAO hybrid 4D-EnVAR DAS. The panels depict surface pressure tendency (hPa/day) on 25 June 2017 from the C720 hybrid DAS (top), and the C180 ensemble mean (bottom).



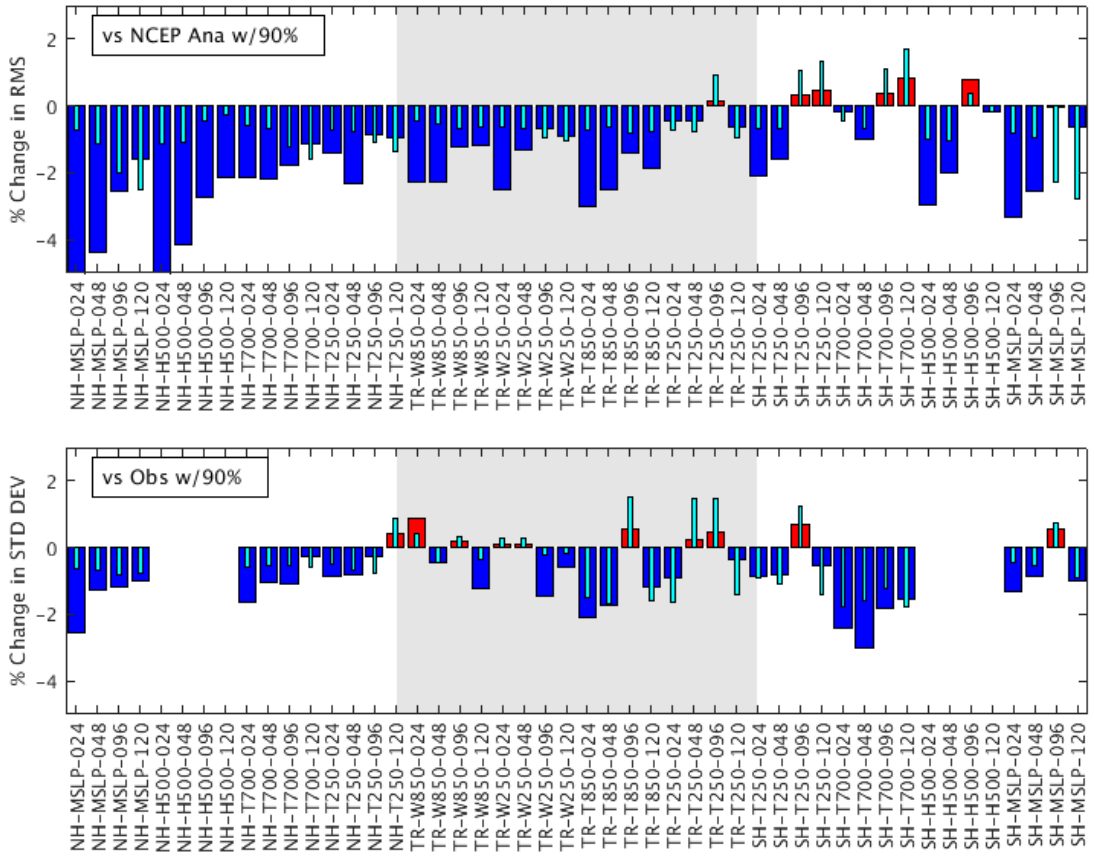
739 FIG. 16. Schematic of the 4D Incremental Analysis Update (4DIAU) used in the GMAO Hybrid 4D-EnVar
740 DAS.



741 FIG. 17. Stability curves associated with the 4DIAU hourly Fixed Analysis (solid dark blue), 4DIAU 3-
742 hourly Fixed Increment with Digital Filter (dashed red), and the 4DIAU 1-hourly Fixed Increment (light blue)
743 with (dashed) and without (solid) Digital Filter.



744 FIG. 18. Similar to Fig. 13, but now fully cycled DAS experiments configured as: FP-like (black curves),
745 3-hourly 4DIAU with digital filter modulation (red curves), and traditional IAU-based 3D-Var (green curves).
746 Lower panel differences are with respect to the FP-like configuration.



747 FIG. 19. Percentage change in forecast scores of comparison between 3-hourly 4DIAU-DF and FP-like con-
748 figurations. Scores are shown for selective quantities, when verified against NCEP analyses (top) and against
749 observations (bottom). Negative/positive values (blue/red bars) indicate improvement/deterioration with respect
750 to FP-like settings; thin cyan bars indicate 90% statistical significance: red and blue bars larger, in magnitude,
751 than thin bars, are statistically significant. Scores cover October-November 2016.